NBER WORKING PAPERS SERIES

THE DEMAND FOR AND RETURN TO EDUCATION WHEN EDUCATION OUTCOMES ARE UNCERTAIN

Joseph G. Altonji

Working Paper No. 3714

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1991

This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #3714
May 1991

THE DEMAND FOR AND RETURN TO EDUCATION WHEN EDUCATION OUTCOMES ARE UNCERTAIN

ABSTRACT

The vast literature on human capital and earnings assumes that individuals know in advance that they will complete a particular program of schooling. This paper treats education as a sequential choice that is made under uncertainty. A simple two period structural model is used to explore the effects of ability, high school preparation, preferences for schooling, the borrowing rate, and ex post payoffs to college on the probability of various post secondary college outcomes and the ex ante return to starting college. The model provides the basis for a simple empirical method of accounting for uncertainty about educational outcomes and for nonlinearity in the relationship between years of education and earnings when estimating the expected return to the first year of college. I present estimates of the effects of gender, aptitude, high school curriculum, family background characteristics, and other variables on the expected return to starting college.

Joseph G. Altonji
Department of Economics
Northwestern University
Evanston, IL 60208
and
NBER

I. Introduction and Summary

Most of the enormous empirical literature on human capital and earnings that has grown out of the work of Jacob Mincer (1958, 1962, 1974), Becker (1975), and other pioneers of the human capital approach to income distribution and earnings abstracts from uncertainty about whether a program of schooling will be completed. An individual chooses the education level that maximizes the present discounted value of wealth given her borrowing rate and the effect of education on earnings, completes her education, and receives the expected return in the market place associated with the chosen schooling level. A rich literature on the returns to education attempts to account for various biases that may arise if education choices are systematically related to other factors (such as aptitude) that influence earnings.[1] But this literature also views the individual as able to choose a future level of education with no uncertainty about actually completing the level. Furthermore, most of the literature on returns to education does not address the fact that there are large differences in earnings by field of study or the fact that choice of college major often changes during college.[2]

This paper examines some implications of the view that educational decisions are made under uncertainty. I analyze a simple structural model of education choice and implement a method for accounting for uncertainty about educational outcomes and for nonlinearity in the relationship between years of education and earnings when estimating the expected return to a year of school. I provide estimates of the effects of aptitude, high school curriculum, and family background characteristics and other variables on the expected return.

The work is motivated by three facts that suggest that uncertainty may have important consequences for the demand for education and for measurement of the returns to education. First, many individuals attend but never finish college even though they report that they plan to complete college or graduate work at the time that they begin college. In the National Longitudinal Survey of

---

1. See for example, Griliches (1977), Willis and Rosen (1978), and the surveys by Rosen (1978) and Willis (1986).

2 Among the recent studies that do examine differences in earnings by college major are Berger (1987), and Paglin and Ruffolo (1990), Brown and Corcoran (1990), and Bamberger (1986). As noted below, Bamberger's model of choice of major explicitly takes into account the fact that individuals who start college in a particular major may switch to another major or drop out of school.

the High School Class of 1972 (NLS72) sample, 89 percent of high school seniors who plan to complete college or graduate school start college but only 58.1 have completed college by 1979.[3] Comparisons of plans with actual majors indicate that individuals are also quite uncertain about what they will major in during college.

Second, for some demographic groups the returns to education are nonlinear, with much of the return associated with completing high school or completing college. For young men I find that attending college for less than two years or for 2 or more years but not receiving a degree changes the log wage by -.0076 and .057 respectively, while obtaining a degree raises the log wage by .179. Nonlinearities are much less important for women in the NLS72 and undoubtedly vary across time.[4] However, the general point is that nonlinearities in the returns may produce substantial differences between ex ante and ex post returns to the first year or two of college. As Weisbrod (1962) pointed out, the return to the first year of college is not the earnings differential between individuals with 12 and 13 years of schooling who are the same in other dimensions that affect earnings. Rather, the return is the difference between the earnings of the person who stops at 12 and the expected earnings net of education costs of a person who attends the first year of college, where the expectation is taken across earnings associated with 14 years, 15 years, 16 years and higher education levels weighted by the probability that the individual who has completed year 13 will successfully complete those higher levels. The coefficients on years of education and on interaction terms involving years of education and school or background characteristics in a conventional earnings regression may provide misleading estimates of both (1) the value <u>ex ante</u> of the additional year of schooling and (2) the effect of school characteristics and high school curriculum on that value. This is because such regressions condition on the ultimate educational outcome, which is uncertain at the time

---

3 This result is for a sample of 9,032 for whom valid data on educational attainment in 1979 and educational plans as of senior year of high school are available. The small fraction does not appear to reflect unreliability of the responses to the question about education plans: 93.5 percent of the people who obtained a college or graduate degree indicated in 1972 that they planned to obtain a college or graduate degree.

4. Many studies find evidence of a college diploma effect or of higher returns to the latter years of college than to the early years. See for example, Hungerford and Solon (1987) and Card and Krueger (1990, Figure 2).

that the decision about whether to pursue an additional year of schooling is made. Variables that affect the probability that an individual will actually complete the program may play an important role in the decision to start college.

Third, there are large differences across fields of specialization in the earnings differential between college and high school graduates. The estimates in Table 1 indicate, for example, that engineers receive high returns in the labor market. The differences across fields are reflected in surveys of the salaries of college graduates and in studies using other micro data sets.[5] Specificity of the knowledge acquired in particular fields of study, such as mathematics or English, and the importance of pre-requisites in some fields, particularly mathematics and the sciences, may mean that course selections in high school or the early years of college condition the options available later in much the same way that occupation specific on the job training conditions future employment options.[6] Decisions to change fields because of new information about preferences for particularly types of work and study or poor performance, or other factors may be very costly. As Bamberger (1986) points out, to the extent that individuals are uncertain about the probability that they will be able to and will want to complete a program in a particular field, they must take into account the alternative options that starting a particular program of study might lead to.

There are a few studies in addition to those mentioned above that have examined the implications of uncertainty about education outcomes. Manski (1989) uses a series of models of education as a sequential decision to analyze the properties of policies designed to reduce the drop out rate from post

---

5. See Berger (1988), who uses the National Longitudinal Survey of Labor Market Experience Youth Cohort. Paglin and Ruffolo (1990) cite several surveys of starting salaries of college graduates. They find a strong relationship between the salaries in different fields and the average Math GRE scores of persons who took the GRE and majored in the field. They attribute much of the difference in wages across fields to the market returns to the math aptitude of those who go into the various fields, but do not examine whether the returns to math and verbal aptitude depend upon the major one graduates in. The present paper does not focus on the sources of the differences in returns across fields, but it is worth noting that the estimates below indicate that there are large differences in the ex post payoffs to various majors even after one conditions on the aptitude and achievement scores and other high school and family background variables.

6. See Shaw (1987) and Sicherman and Galor's (1990) discussion of occupations.

secondary school.[7] Similar ideas arise the occupational choice models of Miller (1984) and Shaw (1987). Bamburger (1986) explicitly accounts for the fact that individuals must take account of the probabilities of having to switch majors or drop out of school and the resulting payoffs when they consider the present value of future earnings associated with an initial choice of major. He formulates and estimates a model of the choice of major in which students base decisions about major on both the expected returns if they successfully complete the major and the probability that they will do so given past educational choices and outcomes. However, the theoretical and empirical literature treating education as a sequential decision under uncertainty is very limited, and uncertainty about completing school has been ignored in empirical studies of the returns to education.

Section II provides a simple 2 period structural model that captures most aspects of the schooling decision problem mentioned above. I use to the model to illustrate the effects of aptitude, high school preparation, tastes for schooling, and the ex post returns to various college degrees on the ex ante internal rate of return to starting college and on the probabilities of various post secondary outcomes.

Section III presents an empirically estimable reduced form version of an equation for the ex ante rate of return to starting college presented in section II. An implication of the structural model in section 2 is that prior to choosing whether to start college each individual faces a set of probabilities that she will ultimately complete a particular level of education in a particular field conditional upon starting college. There is also a set of market payoffs to completing an education program of a particular length in a particular field. Personal characteristics affect the expected rate of return to starting college both by altering the market payoffs associated with completing particular postsecondary programs of study and by altering the probabilities that the individual will complete the programs.

In Section IV I discuss the data and the specifications of reduced form equations for the probability (conditional upon starting college) of 18 education outcomes. I also discuss the variables that are included in the wage equation that is used to estimate the ex post payoffs to various education outcomes. Section V presents estimates of the expected internal rate of return

---

7. See also Comay et al (1973).

to starting college for various groups. In section VI I discuss the sensitivity of the results to treatment of unobserved heterogeneity in the payoffs to various education outcomes and to assumptions about the information set available to high school seniors.

The main empirical results are as follows. First, the ex ante return to starting college evaluated at the sample mean for male high school graduates is 2.8 percent in the base case considered below, while the ex post return is actually slightly negative.[8] The ex post and ex ante returns for female high school graduates are closer ((8.5) versus (7.4)), and are much larger than the return for men. The ex ante returns for both men and women are substantially higher if one allows the ex post payoff to education to increase with labor market experience.

Second, the differences between men and women in the ex post payoffs to various post secondary outcomes explain most of the gender difference in returns. (Women who do not attend college earn much less than comparable men.) Differences in the characteristics of men and women do not make much difference. However, gender differences in the equations for education outcomes tend to raise the return to college for men relative to women.

Third, high aptitude individuals have a substantially higher ex ante return to starting college than low aptitude individuals. Aptitude raises the return for men both by increasing the ex post payoffs to college and by favorably altering the probabilities of the various education outcomes conditional on starting college. For women the link between aptitude and the ex ante return to education is dominated by the effect of aptitude on the ex post payoffs to college.

Fourth, an academic high school curriculum and a favorable family background raises the ex ante returns in the case of young men but makes only a small difference for young women.

Fifth, those who start college are estimated to have a substantially higher ex ante internal rate of return to doing so than those who do not.

In section VII I close the paper with a research agenda.

II. Education as a Sequential Choice

This section uses a simple two period model of education choice under

---

8. The estimate of 2.8 refers to the base case in which the returns to education do not depend upon labor market experience. I report estimates 4.1 and 7.3 under alternative assumptions about the size of the interactions.

uncertainty to flesh out some of the ideas in the introduction and to provide a theoretical foundation for the reduced form equations for education outcomes used in the empirical analysis of ex ante returns to education. In section II.1 I set up the model. In Section II.2 and Appendix 1 and 2 I discuss comparative statics concerning the value of starting college, the ex ante probability of various education outcomes, and the ex ante internal rate of return to starting college.

II.1 The Model

In period 0 an individual decides whether to work or to begin college in either math/science or the humanities. The field of study influences the mix of science and humanities courses taken. Knowledge in math/science and humanities at the end of a year of college depends upon the field chosen, aptitude, the stock of knowledge at the start of the year, and an error vector. The monetary return to the second year of college is degree specific and is 0 if the individual does not attain the minimum (field specific) knowledge requirements for a degree in either field. Furthermore, a person who attends college discovers whether she dislikes college relative to working and her relative preferences for math/science and the humanities. At the end of the first period a student may change her field of study or drop out of school. Her decision reflects new information about preferences, the probability that she will be able to complete a degree in math/science or humanities given her stock of knowledge after the first year of college, and the payoffs associated with the different education outcomes. If she attends school a second year, she finds out whether she completed degree requirements at the end of the year. She then goes to work at the appropriate wage.

I now turn to the details of the model. I first discuss the stock of knowledge, ability, and fields of study and then earnings and preferences.

The Stock of Knowledge: In primary school, secondary school, and postsecondary school individuals acquire knowledge in field m and field h. Let knowledge at the end of school year s be denoted by K_s, $K_s=(K_{ms}, K_{hs})$, where K_{ms} is math and science knowledge, K_{hs} is humanities/social sciences knowledge at the end of s, and s = 0, 1, and 2 refer to high school, the first year of college, and the second (final) year of college respectively.

Ability: At the end of high school individuals differ in two dimensions of cognitive ability that have value in human capital investment and in production. Let $A=(A_m, A_h)$ where A_m is math/science ability and A_h is verbal ability. Presumably A_m is more important for the production of K_{ms} and A_h is more important for the production of K_{hs}.

Fields of Study: A postsecondary school program is a field specific function relating K_{s+1} to K_s, A, and a stochastic component. The field of study in year s is denoted by c_s, c=m or h. Field m is math, science and engineering. Field h denotes humanities and social sciences. The distribution of K_{t+1} is stochastically increasing in K_s and in A regardless of what field is chosen.

A specific example of such a function is

$$(1) \qquad K_{s+1} = \Pi_{cs} K_s + \pi_{cs} A + \epsilon_s \quad .,$$

The matrices Π_{fs} and π_{fs} capture the fact that the particular courses associated with a field of study influence the evolution of knowledge. The first term in the equation captures the idea that K_{s+1} will depend not only on the program of study but also on what one already knows. In fields in which prerequisites are important, such as math and science, one would expect Π_{cs} to place a large weight on the component of K_s that is a major part of the field. The second term in the equation captures the idea that math and verbal ability will influence how much an individual learns from a given program of study. The error vector ϵ_s captures the influence of particular teachers and courses and unforcastable individual specific shocks (such as illness, emotional problems) that affect how much the student learns in the year.[9] The aptitudes A_m and A_h are fixed and known to the student when s=0, so ϵ affects K but not A.

Degree Requirements: A college degree in field h or in m requires that a field specific function of knowledge K_{2m}, K_{2h} exceed a threshold degree requirement. The level of K_m (K_h) is presumably more important in field m (h) than in field h (m). Given that the links between K_0 and K_1 and between K_1 and K_2 are stochastic, there is uncertainty about whether an individual who sets out to complete a degree in field c will succeed. The probability g_{2c} that a person who is studying in field c in the second year will complete the degree requirements depends upon K_{1m}, K_{1h}, A_m, and A_v according to

$$(2a) \quad g_{2m} = G_{2m}(K_{1m}, K_{1h}, A_m, A_h)$$

$$(2b) \quad g_{2h} = G_{2h}(K_{1m}, K_{1h}, A_m, A_h),$$

where both functions are strictly increasing in all arguments. The graduation

[9] One would not expect ϵ to enter additively or to be independent of field.

probabilities g_{2m} and g_{2h} have a distribution conditional on K_0, A, and the choice of field in the first year of school. (This is because K_1 is random). Let $F_{hm}(g_{2h}, g_{2m} \mid K_0, A, c_1)$ be the joint distribution function of g_{2m} and g_{2h} and let $F_h(g_{2h} \mid K_0, A, c_1)$ be the marginal distribution of g_{2h}. I assume that both joint distribution and marginal distribution have a monotone likelihood ratio property in the elements of K_0 and A. In the bivariate case this means that

$$dF_{hm}(x_1, x_2 \mid K_0', A', c_1)/dF_{hm}(x_1, x_2 \mid K_0'', A'', c_1),$$

is strictly increasing in both x_1 and x_2 over the range of g_{2h} and g_{2m} if $(K_0', A') > (K_0'', A'')$, where dF_{hm} is the density corresponding to F_{hm}. The monotone likelihood ratio property implies first order stochastic dominance.

I assume that the programs of study and the requirements for field m and field h are sufficiently different that students who choose to study in field m (h) in the second period have a negligible chance of completing the degree requirements in h (m).[10]

<u>Earnings</u>: The discounted present value of earnings is Y_0 for persons who enter the labor market after high school, $Y_1 = Y_0(1 + r_1)/(1 + R)$ for individuals who leave school after 1 year of college, $Y_1/(1 + R)$ for persons who attend college a second year but fail to get a degree, $Y_{2m} = Y_0(1 + r_{2m})(1 + R)^{-2}$ for persons with a math/science degree, and $Y_{2h} = Y_0(1 + r_{2h})(1 + R)^{-2}$ for persons with a humanities degree. R is the discount rate. I assume

$$Y_{2m}, Y_{2h} > Y_1, Y_0.$$

These inequalities guarantee that persons who are indifferent between school and work will choose to complete college if they are certain of being able to meet the requirements. I assume the experience profile of wages is independent of years in school and field of study.[11] I also assume here that no

---

10 This assumption is a statement about the values of the parameters of the knowledge accumulation equation function, including the degree of randomness, about the levels of degree requirements in the math/science and in humanities.

11. One may regard the earnings parameters r_1, r_{2h}, and r_{2m} as conditional on A and K_0. The comparative statics results reported below for A and K_0 below hold the earnings parameters constant while varying A and K_0. This focuses attention on the fact that these variables influence the value of starting college and the ex ante rate of return even if they do not affect the ex post payoffs to education. One may easily calculate the total effects of A and K_0 by allowing the earnings parameters to vary with them. See also the discussion in footnote 33. In the empirical work I allow the link between education and earnings to depend on ability and experience. One might also wish to allow earnings to depend in a more continuous way on the stock of knowledge.

new information about Y_{2m}, Y_{2h}, and Y_1 arrives until education decisions are made. In Appendix 3 I analyze a variant of the model in which new information about Y_{2m} and Y_{2h} arrives after the first year of college, and the empirical analysis below allows for the possibility of new information.

Preferences: Utility depends on the present value of income, a taste parameter j summarizing nonpecuniary preferences for education program and the job types it leads to, and the type of job and education program an individual chooses. There are three types of people, with j = 0, 1, and 2. Type 0 people dislike the school and white collar jobs sufficiently that they receive a minimal increase in utility to completing a degree even though they would earn higher wages if they complete college and take a white collar job.

Type 1 people are indifferent between spending time at work or at school in the humanities. They are also indifferent between jobs that do and do not require postsecondary education. However, type 1 individuals hate spending time in field m in school or on the job. Their dislike is sufficiently large relative to $Y_{2m} - Y_{2h}$ and $Y_{2m} - Y_1$ that they never choose field m in year 2.

Type 2 individuals are indifferent between the type of jobs and between time spent in school and at work.

At the end of high school the probability that a given individual with characteristics X is type j is equal to $\theta_j(X)$ where j=1, 2, or 3. Individuals learn their preferences after the first year college.[12]

II.2 The Return to College and the Probability of Completing College

I now use the model to examine the effects of various factors on the

---

rather than simply upon the thresholds for a degree. This would complicate the theoretical analysis because one would not be able to summarize the effect of K_1 and A on V_2 as operating through g_2. As noted in the conclusion, one could modify the empirical analysis of ex ante rates of return to allow earnings to depend upon grades and courses using data from postsecondary school transcripts.

12. I am assuming that all individuals have no nonpecuniary preference between spending a year working, spending the first year of college in math, and spending the first year of college in humanities. This could easily be relaxed. To keep the dynamic program problem simple I have ruled out the possibility that those who choose not to go to college because they think they are type 0 individuals revise their prior beliefs and then choose to attend school. There are no conceptual difficulties in relaxing this assumption. One might also wish to assume that the new information about preferences that an individual acquires depends upon the field that he tries, as in the matching models such as Miller (1984). This would dramatically complicate the decision problem because the choice of field also conditions accumulation of K.

value of starting college, the probability of various education outcomes conditional on starting college, and the ex ante return to starting college.

At the end of period of 0 the individual must choose between going to work, attending the first year of college in field m, and attending the first year of college in field h. The value of going to work is simply Y_0. The value of attending the first year of college in field m or field h is

$$(3) \quad V_1(K_0, A, \theta, m_1) = \sum_{j=0}^{2} \theta_j \; E(V_2(g_2, j) | K_0, A, m_1) ,$$

$$(4) \quad V_1(K_0, A, \theta, h_1) = \sum_{j=0}^{2} \theta_j \; E(V_2(g_2, j) | K_0, A, h_1) ,$$

where the expectation E is taken over the distribution $F_{hm}(g_2 | K_0, A, c_1)$ of g_2 conditional on K_0, A, and the choice of field c in the first period and $V_2(g_2, j)$ is the value having attended school for 1 year and obtained graduation probabilities g_2 for a person who is type j.[13]

Type 0 persons drop out school and receive Y_1, so

$$(5) \quad V_2(g_2, 0) = Y_1.$$

For type 1 individuals (who hate m), the value function is the max of the return to staying in college and studying humanities and leaving school and receiving Y_1:

$$(6) \quad V_2(g_2, 1) = \text{Max}\{[g_{2h} Y_{2h} + (1 - g_{2h}) Y_1/(1 + R)], Y_1\} .$$

Type 2 individuals either stay in school and major in h, stay in school and major in m, or drop out and receive Y_1, so

$$(7) \quad V_2(g_2, 2) = \text{Max} \{[g_{2h} Y_{2h} + [1 - g_{2h}] Y_1/(1 + R)], \; g_{2m} Y_{2m} + [1 - g_{2m}] Y_1/(1 + R), Y_1\}.$$

An individual starts college if

$$(8) \quad \text{Max}\{V_1(K_0, A, \theta(X), h_1), V_1(K_0, A, \theta(X), m_1)\} > Y_0.$$

A person starts college in field h if (8) holds and in addition

---

[13] Keep in mind that $g_2 = (g_{2h}, g_{2m})$ is determined by K_1 and A in accordance with equations (2) and (3) and that F_{hm} is determined by the equation of motion for knowledge and graduation requirements.

(9) $V_1(K_0, A, \theta(X), h_1) > V_1(K_0, A, \theta(X), m_1)$.

In proposition 1 I summarize a few properties of the value of starting college. (See Appendix 1 for proofs.)

Proposition 1: The value of starting college relative to Y_0 is increasing in K_0, A, $\theta_2/(\theta_1 + \theta_2)$, Y_1/Y_0, Y_{2h}/Y_0, and Y_{2m}/Y_0 and decreasing in θ_0 and R.

The intuition for these results is simple. K_0 and A shift out the distribution of the probability that the individual will be able to satisfy graduation requirements in math, science, or both. An increase in the ex post payoffs to the completing 1 year of college (Y_1) and to completing two years of college with a degree in math (Y_{2m}) or humanities (Y_{2h}) also raises the value of starting college. θ_0 raises the dropout probability, lowering the return. An increase in $\theta_2/(\theta_1 + \theta_2)$ (holding θ_0 constant) raises the odds that a math degree will be a viable option in the second period. This raises the return to starting college. An increase in the discount rate R lowers the return to starting college even when Y_1/Y_0, Y_{2h}/Y_0, and Y_{2m}/Y_0 are held constant because it lowers the present value of earnings associated with attending college a second year but failing to graduate. Of course, most of the negative effect of R on the value of starting college is due to the fact that an increase in R lowers Y_1/Y_0, Y_{2h}/Y_0, and Y_{2m}/Y_0.

Due to space limitations, I do not analyze in detail the determinants of the value of starting college in m relative to h. The relative value depends in part upon how the field in year 1 and the relative values of K_{0h}, K_{0m}, A_h and A_m influence the distribution of K_1 and as a consequence the distribution F_{hm} of g_{2h} and g_{2m} in year 2. The return to m relative to h is a negative function $\theta_1/(\theta_1 + \theta_2)$ and is likely to be a negative function of Y_{2h}/Y_{2m}, K_{0h}/K_{0m} and A_m/A_m.

Finally, it is interesting to note that sensitivity of the value of college to knowledge and ability depends upon preferences. The variables K_0 and A are relevant to the decision to start college because they increase the probability that the individual could actually meet graduation requirements if she were to choose to try. However, the expected value of the option is diminished if the individual believes that she will turn out to be type 0 and will choose to drop out anyway. If θ_0 is close to 1, then individual does not expect to complete college, and the contribution of K_0 or A to the knowledge accumulation process in college is of little value.

Education Outcome Probabilities and the Ex Ante Return to Starting College

The internal rate of return ρ_c to starting college in field c is the value of ρ_c that solves

$$(10)\ 1 = P_{1c1.}(1 + r_1)/(1+\rho_c) + P_{1c2.}(1 + r_1)/(1+\rho_c)^2 + P_{1c2h}(1 + r_{2h})/(1+\rho_c)^2 + (P_{1c2m})(1 + r_{2c})/(1+\rho_c)^2,$$

where

$P_{1c1.}$ = probability of leaving college after the first year

$P_{1c2.}$ = probability of spending 2 years in college but failing to complete a degree

P_{1c2h} = probability of completing a degree in humanities

P_{1c2m} = probability of completing a degree in math

c = h or m.

In general, the ex ante internal rate of return to starting college is different from the ex post return to the first year. In particular, even $r_1 = 0$, the ex ante internal rate of return to starting college is positive provided that r_{2h} and r_{2m} are positive and there is some possibility of completing college.

In appendix 2 I derive expressions for the education outcome probabilities as a function of θ, K_0, A, Y_0, Y_1, Y_{2m} and Y_{2h}, and the interest rate used to discount earnings. I summarize a few properties of these probabilities in the following propositions. (See Appendix 2 for proofs.)

Proposition 3a. The probability $P_{1c1.}$ of dropping out of college after one year is a positive function of r_1 and θ_0. It is a negative function of $\theta_2/(\theta_1 + \theta_2)$, r_{2h}, r_{2m}, K_0, A, and $1/(1 + R)$.

The taste probability θ_0 increases $P_{1c1.}$ because it increases the odds that the person will find out that she dislikes college and dislikes the types of jobs that reward a college degree and thus decide to drop out. $\theta_2/(\theta_1 + \theta_2)$ is the conditional probability that the person can accept both engineering and humanities, rather than simply humanities. The possibility of majoring in either humanities or mathematics raises the expected value of attempting college. The higher r_{2h} and r_{2m} (holding r_1 constant) and the lower R, the lower the critical value g_2 (the pass probability) for the individual to remain in college after the first year. Increases in K_0 and A increase the probability that g_2 will be high enough to warrant staying in college, lowering the dropout rate after the first year.

Proposition 3b. The probability $P_{1c2.}$ of attending college for a second year and failing to get a degree is negatively related to θ_0 and r_1. The effects of K_0, A, $\theta_1/(\theta_1 + \theta_2)$ are ambiguous. r_{2m} has a positive effect when $r_{2m} > r_{2h}$. r_{2m} and r_{2h} have a positive effect when $r_{2m}=r_{2h}$.

The negative effect of θ_0 on P_{1c2} is obvious given that all type 0 individuals leave school after 1 year. The ambiguity underlying K_0 and A is interesting. Increases in these variables shift out the distribution of the vector of graduation probabilities g_2. This shift increases the likelihood that the realization of g_{2m} or g_{2h} will be large enough to justify attending college for a second year. This will tend to increase $P_{1c2.}$, assuming that the probability of graduating conditional on attending the second year is unchanged. However, the shift will also increase the conditional graduation probability. This will lower $P_{1c2.}$.

The ambiguity in the effect of $\theta_2/(\theta_1 + \theta_2)$ is also interesting. As noted in proposition 3a, the probability of dropping out after the first year is negatively related to $\theta_2/(\theta_1 + \theta_2)$, which means that the fraction of students who attend a second year of college rises. This will tend to increase $P_{1c2.}$ as well as P_{1c2m} and P_{1c2h}. Furthermore, if $r_{2m} > r_{2h}$, some individuals who would have chosen humanities if they were type 1 will choose math even though $g_{2m} < g_{2h}$. This will tend to lower the probability of completing degree requirements conditional on attempting the second year and raise $P_{1c2.}$ as well. On the other hand, the distribution of the maximum of g_{2m} and g_{2h} stochastically dominates the distribution of g_{2h}, and this will tend to increase the conditional graduation probability for type 2 individuals relative to type 1 individuals.[14] The intuition here is that individuals who are only marginal in the humanities but are very good at math/science will have a higher graduation probability if their preferences are such that math/science is a viable option.

Proposition 3.c The graduation probability $P_{1c2h} + P_{1c2m}$ is decreasing in θ_0, (1 + R), and r_1. It is increasing in r_{2h} if $r_{2h} \leq r_{2m}$ and is increasing in r_{2m} if $r_{2m} \geq r_{2h}$. It is increasing in r_{2h} and r_{2m} holding $r_{2h} - r_{2m}$ constant. If $r_{2m}=r_{2h}$, the graduation probability is increasing in K_0, A, and $\theta_2/(\theta_1 + \theta_2)$.

The effect of θ_0, (1 + R), r_{2m} and r_{2h} are obvious. An increase in r_{2h} or r_{2m} will lower $P_{1c1.}$ by raising the ex post payoff to those who complete college. However, the effect of r_{2h} on $P_{1c2h} + P_{1c2m}$ is ambiguous without further assumptions because the increase may lure students with high g_{2m} to take

---

[14] If $r_{2m} = r_{2h}$, the conditional graduation probability is definitely higher for type 2 than for type 1 individuals.

a chance on graduating in humanities, raising the failure rate. The possibility can be ruled out if $r_{2h} < r_{2m}$, in which case students choose humanities over math only if $g_{2h} > g_{2m}$. The same argument applies to r_{2m}.

Increases in K_0, A, and $\theta_2/(\theta_1 + \theta_2)$ lower $P_{1c1.}$. However, they could lower the graduation probability by increasing the fraction of type 2 individuals who opt for the high risk, high return major. This possibility is ruled out if r_{2m}-r_{2h}. Below I find that increases in K_0 and A (as measured high school courses, grades, and tests) increase the graduation probability.[15]

Comparative Statics Analysis of the Expected Internal Rate of Return

One can analyze the effects of ability, preferences, and the ex post payoffs to college by substituting the equations (A2.2), (A2.3), (A2.4) and (A2.5) into (10) and performing comparative statics. The comparative statics relating the ρ_c to $P_{1c1.}$, $P_{1c2.}$, P_{1c2h}, P_{1c2m}, r_{2h} and r_{2m} and r_1 are straightforward. Given the assumption Y_{2m} , $Y_{2h} > Y_1$, ρ_c is negatively related to $P_{1c1.}$ holding $P_{1c2.}$ and P_{1c2h}/P_{1c2m} constant, negatively related to $P_{1c2.}$ holding P_{1c2h}/P_{1c2m} constant, and positively related to P_{1c2m}/P_{1c1h} holding $(P_{1c2m} + P_{1c2h})$ constant. ρ_c is increasing in the ex post returns to college.

The analysis of the effects of the fundamental variables A, K_0, r_{2h}, r_{2m}, r_1, and θ is complicated because they influence all of the probabilities. It is tedious to work through the comparative statics analytically. It is easy to do so numerically, and I do so below.

In the present case, however, the special assumptions made about preferences have the implication that the internal rate of return to starting

---

15. It should be keeping with the empirical emphasis in the paper on the ex ante return to education, Propositions 3.a, 3.b and 3.c characterise how a number of factors influence on the probability of various postsecondary education outcomes conditional on starting college. However, by considering K_0, A, θ, to have a distribution in the population, one may use the model to analyse the determinants of the probability of obtaining a college degree in humanities, in math without conditioning on starting college. The implications may be quite different from a perfect foresight model. For example, with sequential choice under uncertainty, increases in r_{2h} can actually increase the probability that a high school graduate will end up with a degree in engineering provided that $\theta_2/(\theta_1 + \theta_2)$ is not very large and that there is limited specialization in the first year of college. In this case, r_{2h} will have a relatively large effect on the number of persons who attend at least one year of college. This increased flow into the first year of college may dominate the negative effect of r_{2h} on the probability of choosing engineering conditional on attending college one year and being a type 2. (See Appendix 2, equation A2.5) A full analysis of these issues would require a separate paper.

college is monotonically related to the value of starting college. Consequently, Proposition 1 implies that ρ_c is increasing in A, K_0, r_{2h}, r_{2m}, r_1, and $\theta_2/(\theta_1 + \theta_2)$ and decreasing in θ_0.

III. An Econometric Framework for Measuring the Ex Ante Return to Education

In this section I begin by extending equation (10) for the internal rate of return to consider a wider set of post secondary outcomes, including several fields of specialization and advanced degrees. I also provide a specification for earnings associated with the various education outcomes that may be used to compute the ex post payoffs to education that appear in (10).

The expected present value of earnings conditional on attending school level s' + 1 in field c" for a person who has completed s' years of schooling with c' as the most recent field of specialization is

$$(11)\quad PV(s'+1,\ c''|s',c',X,Z,R) = \sum_{s=s'+2} \sum_{c} P(X,Z)_{s'+1c''sc}\, Y(X,Z,R)_{sc} + P(X,Z)_{s'+1,c'',s'+1,c''}\, Y(X,Z,R)_{s'+1,c''}$$

where $Y(X,Z,R)_{sc}$ is the expected present value of earning (conditional on X, Z and the interest rate R) of obtaining schooling level s in field c, $P(X,Z)_{s'+1,c'',s,c}$ is the probability that a person with characteristics X, Z will end up with schooling level s in field c given that they currently are at schooling level s'+1 in field c". As illustrated in the two period model, these education outcome probabilities reflect sequential decisions that are made after each year of schooling based upon information about performance in school, grade and course requirements associated with particular programs of study, the wages associated with different educational outcomes, and preferences for particular fields of study and work. However, to analyze the ex ante return to education, one may work with the reduced form equations instead of the structural equations, which is what I do below.

Assuming that in postsecondary school there is no specialization until after the first year, then the expected present value of earnings for attending the first year of college for a person who has completed high school ($s=0$) and has characteristics X,Z is

$$(12)\quad PV(1,.|0,X,Z,R) = \sum_{s=1} \sum_{c} P(X,Z)_{1.sc}\, Y(X,Z,R)_{sc} + P(X,Z)_{1.1.}\, Y_0(X,Z,R)\ ,$$

where $Y_0(X,Z,R)$ is the expected present value of earnings for persons who leave school after high school. The value of R that equates the right hand side of (12) to $Y_0(X,Z,R)$ is the ex ante internal rate of return ρ to the first year of college.

To estimate ρ and examine its dependence on X and Z one must have estimates of the $P(X,Z)_{1.sc}$ functions and the $Y(X,Z,R)_{sc}$ functions. Unrestricted probit models are used for the probability that the highest education level achieved by an individual who started college is less than two years, more than two years but no degree, a college degree in business, etc. Eighteen mutually exclusive outcomes are considered, as described in the data section below.

To measure the returns associated with the various education outcomes, assume that the log wage of a person with post secondary education level s with field c as the final field of specialization in school and with t years of labor market experience is

$$(13) \quad \ln w_t = XB_1 + ZB_2 + r_{sc}(Z) + \alpha st + \Psi(t) \ ,$$

where $\alpha st + \Psi(t)$ is the experience profile of earnings and where $r_{sc}(Z)$ is the difference in the log wage the individual receives if she chooses to leave school with a high school degree and the log wage she expects to receive if it turns out that she chooses s years of schooling in field c. The value $s=0$ corresponding to a high school degree. Equation (13) assumes that early choices of post secondary field do not affect the log wage conditional on s and the final field c. Note that these choices still may affect ex ante returns because they may alter the probability that one can attend school in year s in field c. Aptitude and Achievement and specialization in high school may be captured by elements of the Z vector.

The wage effect of a particular education sc may depend upon personal characteristics Z for two reasons. First, Z may directly influence the effect of sc on productivity and, as a consequence, wages. People with strong math aptitude might get an extra benefit from an engineering degree. Second, if individuals know in advance that they will obtain additional information about the market payoffs to various education outcomes as they go through school, then variables that are related to the ex ante completion probabilities will be related to the wage that the individuals expect to receive in the event that ex post they choose s and c. Appendix 3 demonstrates that this is true even for variables do not directly alter the effect of a given education level on productivity.[16]

Most of the empirical work assumes that Z is observed by the econometrician. Consequently, the wage equation is inconsistent with wage specifications discussed in Willis and Rosen (1979) and Willis (1986), in which both observed and unobserved personal characteristics shift the percentage effect of education on wage rates.[17] The Willis and Rosen analysis implies a relationship between the unobservables affecting education choice (the P(X,Z) functions) and the unobservables affecting the expected wage associated with a particular education outcome. I relax the assumption of no unobserved heterogeneity in expectations about ex post payoffs with a procedure that assumes that an agent's information about ex post payoffs is reflected in measures of her expectations about education. There is some evidence of selection bias, but it has little effect on the results.[18]

The estimates of the probit models of the education probabilities may also be subject to selection bias, since those who choose not to attend college may have different postsecondary outcome probabilities (conditional on the observables) than those who do. Some bias undoubtedly exists, but it is worth noting that the observables that are used have considerable explanatory power. As shown in Appendix Table A.4, the predicted dropout probabilities are much larger for those who do not attend college than for those who do and are very sensitive to the ability measures, family background, and high school curriculum.

---

[16]. For example, people who know at the time that they start college that they will enjoy the study and practice of engineering also know that they are less likely to switch to another field of study or drop out of college if they find out during senior year of college that they are bad at and will receive a low market payoff to engineering. Consequently, the earnings individuals expect (as of senior year of high school) to receive conditional on ultimately completing of engineering may be negatively correlated with preferences for engineering.

[17] These papers assume that educational attainment is based on a once-and-for-all decision, with certainty about the probability of successfully completing the program chosen.

[18]. Berger (1987) estimates wage equation for several majors using data from the original NLS survey of young men with corrections for selectivity bias based upon a reduced form multinomial logit model of major choice. He reports little evidence of selection bias. However, there are severe difficulties in identifying the effects of selection on wages using conventional approaches. I do not view his results or those of the present paper as decisive on the issue of whether selection bias is important.

As noted below, I have also experiment with adding a vector of measures of college plans to the education outcome models as proxies for unobservables infuencing education choice. These add substantially to the explanatory of the education models but do not have significantly alter the estimated relationship between ρ and ability, family background, high school curriculum, and whether or not one started college.

Let R denote the interest rate at which earnings are discounted and assume for the moment that $\alpha=0$. Then the present value of future wages Y_{sc} is[19]

$$(14)\quad Y_{sc}(X,Z,R) = Y_0(X,Z,R)\exp^{r(Z)_{sc} - Rs},$$

$$(15)\quad Y_0(X,Z,R) = Q\exp^{XB_1 + ZB_2}$$

$$\text{where}\quad Q = \int_{t=0}^{T} \exp^{\Psi(t) - rt}\ .^{20}$$

One may compute ρ by substituting the estimates of the probit models of the education probabilities and equation (15) for $Y_0(X,Z,R)$ and (14) for the various values of s and c into (12). The solution will differ from the ex post return $\exp(r_{1.(Z)})$ to starting college unless $\exp(r(Z)_{sc}/s) = \exp(r(Z)_{1.})$ for all s

19. If the ex post payoff to education is stochastic conditional on the information set Z and the fact that ex post the individual has found it optimal to choose Z, then in principle this randomness should be accounted for in going from the log linear equation (13) to the present value of the wage level Y_{sc}. Suppose that the ex post payoff is $r_{sc}(Z) + \eta_{sc}$ for the various values of sc, where the η_{sc} have mean 0. Then the analysis goes through provided that the η_{sc} for each sc have identical distributions. Intuitively, the effect of randomness in the log of the wage on the expected value of the wage level will be the same for all education outcomes (in percentage terms) if the degree of uncertainty about the payoffs is the same for all outcomes.

It should also be noted that adding a transitory error component to equation (13) leads to similar complications. If the distribution of the transitory wage component is the independent of the schooling outcome, then this randomness changes the present value of earnings associated with the various education outcomes by the same factor of proportionality. However, if the transitory variation is related to the education level and/or if the experience slope of earnings depends upon years of schooling and the variance of the transitory component is related to experience, then complications arise. In principle, one could estimate the uncertainty and modify the internal rate of return calculations, but I have not attempted to do so.

20. Equation (14) assumes that T does not depend on s and c. Mincer (1974) presents evidence that this is a reasonable approximation for s.

greater than 1 and all fields c. Below I compute estimates of ρ for various values of X and Z.[21]

IV. Data and Econometric Specification

IV.1 The Sample

The NLS72 is a Department of Education survey of individuals who were high school seniors during the 1971-1972 academic year. The initial interview was conducted during the Spring of 1972, with followup surveys in 1973, 1974, 1976, and 1979. A subsample was resurveyed in 1986. A subsample of 10306 met various sample selection criteria and had valid data on the variables used in the education analysis.[22] The equations for education outcomes are conditional on starting college, which reduces the sample to the 6660 individuals who have at least some college.

The cross section-time series of observations for each individual used in the wage analysis was created using information on earnings divided by hours for 1977, 1978, and 1979, and information on the wage at the beginning and end of each job held between 1980 and 1986 up to a maximum of the four most recent jobs.[23] An observation for 1977 is included if (1) the individual was not a full time student in October 1976 nor October 1977, (2) hours worked in 1977 was greater than 1,040, <u>and</u> (3) the 1977 real wage was between \$.50 and \$75 in 1967 dollars. Observations for 1978 and 1979 were included if they met the corresponding three criteria for 1978 and 1979 respectively. Data for begin and end job dates (1980-86) were included if (1) the hours worked in the appropriate

---

21. The present value formula is modified slightly if wage growth rates depend upon the education level, which I consider in the empirical work.

22. The NLS72 contains data on 22,652 people, 12,841 of whom were resurveyed in 1986. I restrict the sample first to the 16,683 individuals from the schools that participated in the base year survey, then to the 15,680 for whom high school test information is available, and then to the 12980 individuals who were surveyed in each of the 1973, 1974, 1976 and 1979 followups. Information from the 1986 follow-up was then added, and only those 7358 persons who were in the earlier 12980 sample were included. The sample of 12,890 from the 1972-1979 surveys forms the basis for the analysis of education outcomes. Of these 10306 had valid data on the variables used in the analysis, which is about 61.8 percent of students in the base year sample and 65 percent of students in the base year for whom test data are available.

23 One potential problem with this sample design is that it is weighted toward persons who have worked for several different employers.

year was greater that 1040, and (2) if the real wage was between $.50 and $75 in 1967 dollars. The wage sample contains 38,595 observations on 9239 individuals.[24] Descriptive statistics are provided in Table A.3.

IV.2 Variable Definitions

A few of the variables require discussion. The high school curriculum measures consist of semester hours in industrial arts, commercial, fine arts, and the 5 main academic subject areas of science, math, social studies, English, and foreign language.[25] I also include dummy variables for whether the student was in the academic track or in the general track.

The education outcomes measures include dummy variables for whether the highest level of education an individual had completed as of 1979 is postsecondary vocational education and no college (VOC79), less than 2 years of college (SOCO1479) and 2 or more years of college but no degree (SOCO1579).[26] Occasionally I use dummy variables for whether an individual has a college degree but not advanced degree (COLL79) and whether the individual has an advanced degree (ADV79). The fields of the college and advanced degrees are aggregated from the 4 digit codes reported in the 1979 survey. The college majors consist of

1. Business (including economics and communications)
2. Engineering and Technical
3. The Physical Sciences
4. Humanities (English, Foreign Languages, and Theology)
5. Social Sciences, including History, Psychology, Legal Studies, Consumer Services, and Area Studies,
6. The Life Sciences and Health Fields
7. Education, including Home Economics and Library Sciences
8. Mathematics, Computer Science
9. Fine Arts

---

[24] The number of observations per person in the wage sample is 4.177 with a standard deviation of 2.34 and a range from 1 to 11. Fifty-two percent of the wage observations come from 1977, 1978, or 1979, 21 percent come from 1985 or 1986, and 27 percent come from 1980, 1981, 1982, 1983, or 1984.

[25] The measures of semester hours refer to courses taken between July 1, 1969 and the date the student will graduate, and so refer to 10th, 11th, and 12th grade for most students. The information was provided by the high schools. The semester hour variables were computed by taking the sum of the semester courses in each subject area, weighting each semester course by the number of hours per week that it met. See Altonji (1988).

[26] I use education as of 1979 rather than education as of 1986 because not all students are included in the 1986 followup and because the sampling probabilities for 1986 depend upon educational outcomes.

10. A miscellaneous category consisting of 7.8 percent of the persons whose highest education level is a 4 year college degree.

The graduate fields consist of Business, including legal studies and communications, Social Sciences and Humanities, Technical Fields (including Engineering, Mathematics, and Computer Science), the Life Sciences and Health, Education (including Home Economics and Library Sciences) and Miscellaneous, including fine arts, consumer services, and missing). One must aggregate to keep the analysis manageable and because of sample size considerations.[27] The college major indicators are coded as 0 if the individual has an advanced degree in any field. High school is the omitted category in the wage equations.

The aptitude and achievement measures consist of predicted SAT_Math and SAT_Verbal scores, high school grades and the student's own assessment (on an inverse 1 to 5 scale) in the Spring of senior year of whether or not she is college material (COL_ABIL).[28]

The vector X_i in the education equations contains (1) dummy variables indicating whether the individual is female, black and/or hispanic, (2) a set of family background characteristics consisting of the levels, squares, and crossproducts of father's years of education and mother's years of education, the log of family income, and a set of variables that measure parental influence on and aspirations for their children, (3) the levels and squares of the predicted SAT_MATH and SAT_VERBAL scores, high school grades, and COL_ABIL, (4) the high school curriculum measures. X_i also includes the level and square of hours per week spent on homework and controls for the size of community and the region of the country in which the individual lived in 1972.

---

[27] In computing ρ is I set s to 1 for SOC01479, 2.75 FOR SOC01579, 4 for all college degrees and 6 for all advanced degrees. I am ignoring variation in expectations about how long it will take to get a degree.

[28]. The NLSHS72 provides standardized scores from each of a battery of six test that were administered as part of the base year survey. SAT Math and SAT Verbal scores are available for about one third of the base year sample. I constructed a composite measure of verbal and mathematics aptitude by regressing SAT Math and SAT Verbal scores against levels and squares of the 6 tests, the cubed values of the reading, mathematics, and the vocabulary tests, a cubic in COL_ABIL, a cubic in high school grades, and dummy variables for race and gender and taking the predicted values for all sample members. I use the equations to compute predicted SAT_Math and SAT_Verbal scores for each sample member. The R2 for the Math SAT and the Verbal SAT are .711 and .760 respectively. I also experimented with using the base year reading, vocabulary and math tests directly and obtained similar fits for the wage and education outcome equations.

The models are estimated on the subsample of young men, the subsample of young women, and the combined sample (with a gender dummy variable included).

Turning to the wage equations, the wage measure is the log of real hourly wage rate in period t.[29] In the most of the specifications the variables Z_i that affect the payoff to the various post secondary education are restricted to the aptitude and achievement measures. I include interactions between SAT_MATH and SAT_VERBAL and (SOCO1479 + SOCO1579), COLL79 and ADV79, between SAT_MATH, SAT_MATH^2, SAT_VERBAL, and SAT_VERBAL^2 and a dummy variable for whether the individual has a college or advanced degree in a technical field, and between SAT_MATH^2 and SAT_VERBAL^2 and (COLL79 + ADV79). The variable COL_ABIL is interacted with [SOCO1479 + SOCO1579] and with [COLL79 + ADV79]. The X_i and Z_i appear separately in the equation, along with experience, experience squared, and a quadratic time trend.

Some of the equations also contain the term α (EXP*s), which is the coefficient α times the product of experience (EXP) and years of schooling s. I do not rely upon the sample to estimate α because the range of labor market experience of the sample is limited, particularly for college graduates. Instead, I set α to three alternative values prior to estimation of the wage equation and present results for each case. The first case is α=0. The second case is α = .0011, which is the estimate obtained using panel data for male heads of household between the ages of 18 and 60 from 1968-1981 from the Panel Study of Income Dynamics.[30] The third value is .005, which is approximately the point estimate one obtains when one estimates α freely in the sample. This value reflects the explosion in the return to education in the 80's (Murphy and Welsh (1988). I take it to be an upper bound for what the Class of 1972 contemplated at the time they were making their decisions.

V. Results

I begin with a brief discussion of the distribution by education outcome

---

29 The use of the hourly wage rate rather than annual earnings provides a crude standarization for differences across majors in typical hours. However, I am not considering differences across jobs in nonpecuniary attributes. The comparisons between men and women ignore complications associated with gender differences in labor force participation and hours worked. In working with wage rates rather than earnings I am also ignoring the return to education that comes from a reduction in the unemployment probability, as analyzed in Mincer (1988).

30. The sample is described in Altonji and Shakotko (1987). The equation is available from the author.

and the rates of return associated with the difference outcomes. I then turn to estimates of the ex ante return (ρ) to starting college for the pooled, men and women samples. I then consider in more detail the contributions of gender differences in (1) personal characteristics, (2) the wage equations, and (3) the education outcome equations to the gender difference in the rate of return to education. Next I discuss the effects of aptitude, high school curriculum, and family background. Finally, I compare estimates of the return to starting college for college going and non-college going samples.

V.1 Summary Statistics on Ex Post Returns and Education Outcomes

Because of the large number of education outcomes and the large number of interaction terms and nonlinearities in the wage equations, it is convenient to begin with some basic information about the probabilities of the various education outcomes conditional on starting college and the ex post returns. I estimated wage models with all interactions terms between Z and the education outcomes excluded and report the estimates of the ex post effects of the various education outcomes on the log wage (relative to high school) in columns 1, 3, and 5 of Table A.1. Separate estimates are reported for the pooled, male, and female samples.

The results show first that the effect of SOCO1479 on the log wage is .0369 for the sample as whole, implying a percentage wage increase of about 3.7 percent. The coefficient is actually negative for men (-.0076). In contrast, the effect of SOCO1479 on the log wage is .0789 for women. Similarly, the percentage return to men for attending college for two or more years but not obtaining a College degrees is 5.9 (5.9=exp(.0575) , while the corresponding value for women is 21.4 percent (21.4= exp(.19455). These differences at the low education levels have a large effects on the estimated ex ante returns because about 65 percent of both men and women do not complete college.

The results in Table A.1 also show that there are large differences in the ex post effects of the various college and advanced degrees on the log wage. For the pooled sample, engineering (.5105), physical sciences (.3367), math and computer science (.4475) and the physical sciences (.3438) are the college degrees with the largest ex post returns. At the other end of the earnings spectrum are humanities (.1676), education (.1588) and fine arts (.1431). There are broad similarities in the relative returns for men and for women but many of the estimated returns for specific college or graduate fields are larger for women than for men, particularly in the case of business, education, and fine arts.

When one restricts the coefficients to be the same across majors and suppresses the aptitude interaction terms, the return (relative to high school) of college and advanced degrees in technical and nontechnical fields may be summarized as follows:

Effects of College and Advanced Degrees on the Log Wage (standard errors in parentheses)

	Pooled	Women	Men
College, Nontechnical	.238	.298	.152
	(.015)	(.019)	(.026)
College, Technical	.441	.404	.394
	(.023)	(.050)	(.038)
Advanced, Nontechnical	.377	.431	.284
	(.033)	(.043)	(.052)
Advanced, Technical	.521	.824	.419
	(.038)	(.050)	(.078)

The results show a large difference in the returns to college and especially to advanced, technical degrees. The coefficient of .824 for women with an advanced technical degree implies an increase of 128 percent over the earnings of high school graduates. However, only 6 women complete advanced degrees in a technical field. The coefficient on FEMALE in the pooled regression is -.162, and so these results indicate only that the percentage differential between men and women narrows with education.

The probabilities of the education outcomes differ substantially between men and women. The sample probabilities (not conditioned on starting college) are reported column 2, 4, and 6 of Table A.1 for the pooled, male, and female samples (respectively). Women are much less likely to major in business/communications, engineering, and the physical sciences than men, and are much more likely to major in education and somewhat more likely to major in the humanities.[31] These results refer to those who do not go on to graduate school,

---

31. Studies of gender differences in choice of college major include Polachek (1978) and Blakemore and Low (1981), England (1982) and Berryman (1983). In this paper I simply examine the implications of differences in college outcomes for the internal rate of return to starting college. In particular, I am not assuming that education choices are made to maximize the present value of future income.

but the gender pattern is consistent with the gender distribution by 4 year college degree of the combined samples of individuals whose highest degree is either college or advanced. (Table A.2) [32]

V.2 Estimates of the Ex Ante Return to Education

Table 1 reports a variety of estimates of ρ. The column labels in the first row of the table indicate whether the sample used to estimate and to evaluate the education outcome probabilities and the wage and education outcome equations consists of men and women (columns 1 and 2) men only (columns 3 and 4) or women only (columns 5 and 6). The titles of each panel provide information about the values of the X and Z variables used to evaluate ρ.

In Panel 1.A, ρ is evaluated at the mean of all variables computed over the sample that includes both persons who started college and those who did not. For the men and women combined, the ex ante return is 5.1 when α=0. (column 1). In Appendix table A.4 I report the probability of various education outcomes and the ex post payoffs, evaluated at the mean for the pooled sample. The low return in part reflects the high dropout probability: The ex post return to attending college for less than two years is 4.1 percent. Columns 2 and 3 present results based on equations estimated on the sample of males and females (respectively) and evaluated at the sample means for male and females (respectively). When α is set to 0, the ex ante return to starting college evaluated at the sample mean for men is 2.8 percent, while Table A.4 shows that the ex post return to attending college for less than two years of college is actually negative (-.61).[31] The negative coefficient gets a large weight in the estimate of ρ., because the ex ante probability of leaving school with less than 2 years of college is .296 (Table A.4). The ex post and ex ante returns for female high school graduates are closer ((8.5) versus (7.4)), and are much larger than the return for men.

---

32. I also examined the links between undergraduate and graduate field for those who go on to graduate school. Education majors and life sciences/health majors are very unlikely to switch fields at the graduate level. 19 of 25 business/communications majors who attend graduate school concentrate in business, law, and communications. 11 of 15 engineers concentrate in a technical area in graduate school. The rest go to business, law, or communications programs. However, social science/services majors are not very concentrated at the graduate level, and humanities majors are less concentrated than the other groups.

31. The estimate of 2.8 refers to the base case in which the returns to education do not depend upon labor market experience. I report estimates 4.1 and 7.3 under alternative assumptions about the size of the interactions.

The ex ante returns for both men and women are substantially higher if one allows the ex post payoff to education to increase with labor market experience.[32]

<u>Sources of the Gender Difference in the Internal Rate of Return</u>

Is the large gender difference in the rate of return due to gender differences in characteristics, gender differences in the education outcomes equations, or gender differences in the market payoffs to education? In Appendix Table A.5 I report the internal rate of return estimate for the 27 possible permutations of (1) the choice of sample used to estimate the wage equation (pooled, men, and women), (2) the choice of sample used to estimate the education outcome equations (pooled, men, and women with at least some college), and (3) the sample used to compute the mean values of characteristics that are plugged in the wage and education equations when computing the internal rate of return. The results show that higher <u>ex post</u> payoffs for women from most of the college majors and, in particular, to completing some college are responsible for most of the difference. The higher payoffs raise the rate of return by about 5 points. (Compare Table A.5, columns 1.b-9b to columns 1.c to 9c).

Given the differences in the college major distribution and given that these differences are not explained by gender differences in personal characteristics, one might expect that differences in the education coefficients tend to lower the rate of return for women. They do, by an amount that is typically about .9 when $\alpha = 0$. (Compare columns 4a-6a to 7a-9a, 4b-6b to 7b-9b, and 4c-6c to 7c-9c.)

Finally, differences in the personal characteristics of men and women do not have a consistent effect on the difference in returns.

---

32. One obtains somewhat different results if one computes the education probabilities as the mean of the probabilities evaluated over the sample distribution of X. For the pooled sample, the estimated probability of College, LT 2 and College, 2+ are .382 and .282 respectively. For men the probabilities are .351 and .301. For women the probabilities are .407 and .275. The probabilities of the college and advanced categories increase relative to those reported in Appendix Table A.4. When the mean education probabilities (as opposed to the probabilities computed at the sample means) are used along the wage equation evaluated at the sample means, the estimated internal rate of return for the case $\alpha=0$ are 5.5 for the full sample, 3.1 for men, and 7.8 for women. In principle one can compute an internal rate of return for each of the 10340 persons in the sample and average the result, but I have not attempted to do so. One would use the micro data to evaluate the wage equation as well as the education equations.

The Effects of Aptitude and Achievement Differences

Panel 1.B of Table 1 shows the difference in the internal rate of return between individuals who are one standard deviation below and one standard deviation above (respectively) the overall sample mean in SAT_MATH, SAT_VERBAL, GRADES and -COL_ABIL (COL_ABIL is on an inverse scale). Using the means and the model estimates for the combined sample, the estimates of ρ when α=0 are 6.4 and 3.6 for the high and low ability cases. The value of ρ for the high and low ability cases for men are 4.2 and 0.5. The corresponding values for women are 8.8 and 5.9. Thus, there is a large difference in the ex ante return to college for the low and high ability cases.

Does ability raise the return by inducing a favorable shift in the education outcomes or by raising the ex post payoffs associated with the various outcomes? For the sample of men with α=0, ρ is 3.8 and 2.2 for the high and low aptitude cases when one varies the aptitude measures in the wage equation but evaluates the education equations at the sample mean. The corresponding numbers are 8.7 and 6.3 in the case of women. These results in combination with Table 1.B suggest that for women the main effect of higher ability is to raise the ex post payoffs to education. For men higher ability increases the return to college both by raising the ex post payoffs to education and by inducing a favorable shift in the education outcomes.[33] The college dropout probability is .9097 for the low ability males and .5195 for the high ability males.

Family Background:

Panel 1.C compares ρ for those with family backgrounds that are favorable to attendance at postsecondary school with those. The results suggest that a favorable family background raises the return to starting college by a substantial amount for young men. (1.4% when α=0). On the other hand, a favorable background is actually associated with a slightly lower return in the case of women. The probability of graduating from college (conditional upon

---

33. With structural equations analogous to those discussed in Section II, one could distinguish among three channels through which ability affects ρ. First, ability alters the distribution of education outcomes holding expected ex post payoffs constant (e.g., it changes the odds of passing school requirements). Second, it changes the mix of education outcomes by altering the ex post payoffs. Third, it alters the ex post payoffs. The reduced form equations for the education outcomes do not hold constant the expected ex post payoffs. Consequently, the estimates of the affect of ability on the distribution of education outcomes lump together the first two channels.

starting) is .402 for an individual with the mean values for the pooled sample and a favorable background and is .135 for those with an unfavorable background. (Table A.4). The figures for men and women are similar. The effect of background on the graduation probability has a bigger effect on the ex ante return for men than women because of the low return to some college for men.

High School Curriculum

One might expect high school curriculum to affect ρ for at least three reasons. First, large numbers of academic courses may increase the ability of students to handle college courses. Second, the composition of courses may affect the relative probabilities of completing particular degrees because of effects on preferences and because knowledge in particular subjects may be a prerequisite to further study. Third, the course variables and dummies for participation in the academic and general tracks may be indicators of tastes for college and family pressure to attend college. The aptitude and achievement measure and family background measures may not fully control for these factors. In addition, curriculum may alter the ex post payoff to various post secondary outcomes, although the specification of the wage equation used does not allow for this.[34]

Panel 1.D reports ρ for individuals who are in the high school track, take one standard deviation more semester hours in science, foreign language, and mathematics, and one standard deviation fewer semester hours in industrial arts and commercial than the sample means for these variables for the full sample. Column 2 reports ρ for individuals are not in the academic track and who are a standard deviation below the sample mean in science, foreign language, and mathematics and a standard deviation above the sample mean in industrial arts and commercial courses. Both columns use the means and the model estimates for the combined sample. When $\alpha=0$ the difference is .7. The difference for the sample means and equation estimates for young men is 1.0. The corresponding difference for women is only .1. Thus, high school curriculum has a modest effect for men but only a small effect for women. The difference in the probability of graduation between the academic and nonacademic cases is .326 for men and .353 in the case of women. (See Appendix Table A.4) The fact that academic curriculum

---

[34] One could use a wage model that allows high school curriculum to effect the return to the various post secondary education outcomes. I have chosen not to add additional interaction terms to the wage equation, in part because the results of Altonji (1988) suggest that the main effects of high school curriculum on wage rates are relatively small.

has a larger effect on the return to starting college for men is due to the fact that men who drop out of college receive a much lower payoff than women.

Differences in Ex Ante Returns for those Who Do and Those Who Do Not Start College

Panel 1.E reports ρ evaluated at the sample means for the college going and non_college going sample. The main result is that ρ is substantially higher for those who actually did start college than those who did not. The ex ante probability for high school seniors of obtaining a college or advanced degree is much higher for those subsequently choose to start college than for those to do not. The differential is proportionately larger for men, reflecting a huge difference in the probability of ultimately dropping out of college between those who do and do not choose to start college. (See Table A.4).

VI. Modifications to the Empirical Framework

In this section I consider the effects of modifying the information set used in forming expectations about the return to college. I also deal with the problem of selectivity bias in the estimates of the wage equation by using education plans and actual years of education as indicators of unobserved differences in the payoff to education.

VI.1 The Information Set as of Senior Year of High School

The results for men indicate an important empirical distinction between the ex ante and ex post returns to education. These results may be misleading if important variables have been excluded from the information set X, Z that individuals are assumed to use in evaluating the probabilities of various educational outcomes when deciding whether to start college. The final education outcome may appear uncertain given the variables used, but may be known to the individual at the time they start college. As a check on this, I added measures of plans (as of senior year) to attend or transfer to a four year college plans and to obtain a 4 year or advanced degree the education outcome equations.

The results are in Panel 1.F of Table 1. For the pooled sample, ρ is between .6 percent higher for those planning to attend college than for those who did not plan to attend college. The differential for men is about 1 percent. For women, the estimated return is actually a bit higher for those who were not planning to attend college. I suspect that the gender difference in the effect of plans is due to the much larger return to some college for women.

VI.2. Adjustments for Unobserved Differences in the Payoff to Education

The estimates of ρ will be biased if I do not include all the variables Z that determine expectations about the payoffs to various levels of education. Partition Z into the subvectors Z_1 and Z_2. Assume both Z_1 and Z_2 are known to the agent at the time that the decision to attend college is made but that the econometrician does not observe Z_2. Since Z_2 influences the payoff to postsecondary education, almost any model of the demand for education implies that it will influence the decision to start college. In this case, the estimated differential between a high school and college graduate conditional on Z_1 will be biased as an estimate of the agent's expectation of the differential given Z_1 and Z_2.

Assume that the expected contribution to lifetime utility of starting college depends upon X and on the vector of payoffs associated with the different educational outcomes conditional on the information set Z_1 and Z_2 and the fact a particular education outcome will be chosen <u>ex post</u>. Z_2 is unobserved. However, assume that there exists a vector of measures ED^e that indicate how much education the individual expects to attain. These expectations depend on X, Z_1, and Z_2. I make the key assumption that these indicators are an <u>exact</u> function of X, Z_1, and Z_2. I also impose linearity, although this could be relaxed. This leads to

(16) $ED^e = X\beta_0 + Z_1 \beta_1 + Z_2 \beta_2.$

Although Z_2 is unobserved, one may solve for

(17) $Z_2 \beta_2 = ED^e - [X\beta_0 + Z_1 \beta_1]$

I approximate the elements of $Z_2 \beta_2$ as the residuals of regressions of the elements of ED^e against X and Z_1. In practice, the vector ED^e consists of dummy variables for the highest level of post secondary education the individual expects to complete (vocational, some college, or college and advanced), a dummy variable for whether the individual will attend a 4 year college, and a dummy variable for whether the individual will start at a two year college and transfer to a four year college or start at a 4 year college.

It is not practical to allow the education payoffs to depend on this many variables, and so I combine them by estimating the equation

(18) $YRSACD79 = X\, G_1 + Z_1\, G_2 + [Z_2\, \beta_2]\, G_3 + \epsilon.$

Finally, I allow the wages associated with stopping school after high school, some college, college, or an advanced degree (respectively) all to depend upon the index $[Z_2\, \hat{\beta}_2]\, G_3$.

There are two important limitations of the approach. First, if the vector ED^e depends upon factors that are unrelated to the expectations about *ex post* payoffs to education in addition to the X variables we control for, then the estimate $[Z_2 \hat{\beta}_2]G_3$ will be a noisy measure of the heterogeneity in the education coefficients in the wage equation. For example, X may not control adequately for differences in tastes for education, and these tastes are presumably reflected in ED^e. Or the elements of ED^e may be noisy indicators of educational expectations. Second, it is unlikely that one may write $r_{sc}(Z_1, Z_2)$ as $r_{sc}(Z_1, [Z_2 \hat{\beta}_2]G_3)$, which I am assuming when I use of interactions between the educational outcomes and $[Z_2 \hat{\beta}_2]G_3$ in the wage equation.

A second, closely related approach is to regress YRSACD79 on X, Z_1, ED^e, and, form $\widehat{\text{YRSACD79}}$, and add interactions between $\widehat{\text{YRSACD79}}$ and SOCO1479, SOC1579, and (COLL79 + ADV79) to the wage equation. The basic idea is that many of the variables that are related to expected educational attainment (given information as of senior year) may also be related to the ex post payoff to education.

<u>Results</u>

The estimates of the wage equation with $[Z_2 \beta_2]G_3$ included suggest that, at least for the pooled sample, the variable has only a small effect on the return to postsecondary education and incorporating the variable does not make much difference for estimates of ρ.

When I use the second approach to obtain coefficients (uncorrected OLS standard errors) of -.0073 (.0068) for $\widehat{\text{YRSACD79}}$, .0097 (.0083) for SOCO1479*$\widehat{\text{YRSACD79}}$, .0165 (.0084) for SOCO1579*$\widehat{\text{YRSACD79}}$, and .053 (.0096) for [COLL79 + ADV79]*$\widehat{\text{YRSACD79}}$. The positive coefficients on $\widehat{\text{YRSACD79}}$ and the interaction terms indicate that higher values of YRSACD79 are associated with higher wage levels regardless of education, and with a higher ex post payoff to education. In the case of the pooled sample with $\alpha=0$ the estimate of ρ is unchanged when one uses the augmented education and wage equations. I have also re-estimated the differences in returns for the college and noncollege going samples and by ability level, family background, curriculum, and college plans using the expanded wage model.[35] The results are basically similar to those reported in the tables.

---

35. In these calculations the values of the variables used to form $\widehat{\text{YRSACD79}}$ reflect the differences in ability, family background, etc. that are under examination.

VII. Directions for Future Research

This paper presents a theoretical and empirical analysis of the demand for and return to education when educational outcomes are uncertain. The simple structural model implies an equation relating the rate of return to starting college to the earnings associated with each of the possible outcomes of starting college (including dropping out after a year, getting a college degree in math, etc.) weighted by the probability of the specific outcome. The ex ante internal rate of return to education depends on factors that affect the odds of completing various college majors, such as preferences for schooling, aptitude and high school achievement variables, as well as on the ex post payoffs to various postsecondary education outcomes. I evaluate the formula for the ex ante internal rate of return using reduced form models for the probabilities of various education outcomes and estimates of the ex post payoffs to them. I estimate the effect of parental background, high school curriculum, academic ability, and gender on the internal rate of return to starting college. The main results of the empirical analysis are summarized in the introduction and seem quite promising.

There is a long research agenda. On the theoretical side, one could easily follow the suggestion in footnote 15 and use the model to analyze the probability of starting college and the supply of college graduates in various fields. It would be useful to relax a number of the special assumptions of the model. It might also be useful to examine the screening/human capital debate from the perspective of the sequential choice model.[36]

On the empirical side, a number of refinements to the model and

---

36. For example, Lang and Kropp (1988) base a test of human capital models versus screening/sorting models on the idea that an increase in the school that, to a first approximation, increases in the compulsory school leaving age should not alter the distribution of education above the level affected by the increase. They are correct in the case of conventional human capital models that assume perfect foresight about education choices. In the model in the paper, however, a law requiring all high school graduates to attend the first year of college would result in an increase in the fraction of high school graduates who graduate from college, because some of the reluctant college attendees would find out after the first year that they are type 1 or type 2 individuals or have higher graduation probabilities than expected given K_0 and A. (In their conclusion Lang and Crop note that "A compulsory schooling law might affect individuals it is does not directly constrain if...forcing persons to go to school longer teaches them that they are benefiting from higher level schooling.") One would have to expand the model to incorporate explicit assumptions about how firms value knowledge and ability before it could be used to address the issue of education as a screen.

On the empirical side, a number of refinements to the model and improvements in the empirical estimation should be implemented before strong conclusions can be drawn about the factors that influence the ex ante return to starting college. The absence of standard errors on the internal rate of return estimates should also be kept in mind in considering the results.[37] One might also wish to experiment with letting the ex post payoff to college depend upon specific courses taken and grades. In principle, one can do this using data from the Post Secondary Transcript Survey of NLS72. However, it would be necessary to add forecasting equations for grades and course counts to the education outcome model. In view of the large shift in the ex post payoff to college in the 80's, more attention should be given to modelling expectations about the earnings associated with various education outcomes. The problem of sample selection bias in the estimates of the wage and education equations remains a potentially serious issue.

Another natural extention would be to distinguish between the ex ante return to high quality and low quality colleges. Differences in dropout probabilities may be more important than differences in ex post payoffs in determining the ex ante return to attending a particular school.

The most interesting and also the most difficult extension would be to expand the structural model of education decisions in Section II into a model that is rich enough to take to the data. One could then replace the reduced form equations relating personal characteristics as of senior year to education outcomes with a structural model of the sequence of education decisions leading to a final education outcome. With transcript data on courses and grades and with survey information on attitudes toward school, one might be able to devise an empirical counterparts to the equations for knowledge and for graduation probabilities and the preference variables in the theoretical model.

---

37. Estimating the standard errors is difficult because the estimate of ρ is a function of hundreds of parameters from the education equations and the wage equation as well as the values of X and Z that enter the wage and education equations. One could estimate the variance matrix of the model parameters and then use a simulation method to compute the standard error of ρ. However, given the number of parameters involved, it would be very difficult to estimate the variance matrix. Or one could apply bootstrap methods to the entire problem. The bootstrap approach would involve drawing random samples from the sample used in the study, estimating all of the model parameters, computing ρ for each of the random samples, and computing the variance of the result. This approach is beyond my computing resources given the large number of different specifications I examine.

ACKNOWLEDGEMENTS

Research support from the National Center on Education and Employment, Teachers College Columbia University, the Spencer foundation and the Center for Urban Affairs and Policy Research at Northwestern University are gratefully acknowledged. I am indebted to Thomas Dunn and especially to Lew Segal and James Spletzer for excellent research assistance. I am grateful to Robert Willis and participants in a seminars at Columbia, NBER, University College London, Northwestern, U.C.S.D., U.C.L.A., U.C. Berkeley, U. of Washington, U. of British Columbia, University of Western Ontario, and University of Chicago for helpful comments. Some of the ideas in this paper were presented in preliminary form in an unpublished paper entitled, "Accounting for Uncertainty When Estimating the Rate of Return to Education" (September 1988). A May 1990 draft was prepared for a Columbia University conference honoring Jacob Mincer. I am solely responsible for the opinions expressed in the study for any errors or omissions.

REFERENCES

Altonji, Joseph G., "The Effects of Family Background and School Characteristics on Education and Labor Market Outcomes" (December 1988)

______________, "The Effects of High School Curriculum on Education and Labor Market Outcomes" (December 1988)

Altonji, J. and R. Shakotko. "Do Wages Rise with Job Seniority?" Review of Economic Studies (July 1987)

Bamberger, Gustavo, "Occupation Choice: The Role of Undergraduate Education", unpublished paper, Graduate School of Business, University of Chicago (October 1986).

Becker, Gary S., Human Capital: A Theoretical and Empirical Analysis, NBER: New York (1975).

Berger, Mark C., "Predicted Future Earnings and Choice of College Major", Industrial and Labor Relations REview", v41 No. 3, April 1988, 418-429.

Berryman, Sue, Who Will Do Science? Ford Foundation, 1983.

Card, David, and Alan Krueger. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." Working Paper No. 265, Industrial Relations Section, Princeton University, (April 1990).

Comay, Y., A. Melnik, and M.A. Pollachek, "The Option Value of Education and the Optimal Path for Investment in Human Capital", International Economic Review 14 No. 2 (JUne 1973) : 421-435.

England, Paula, "The Failure of Human Capital Theory to Explain Occupational Sex Segregation", Journal of Human Resources 17 pp. 358-370.

Griliches, Z. "Estimating the Returns to Schooling: Some Econometric Problems." Econometrica, January 1977, pp. 1-22.

Hungerford, Thomas and Gary Solon, "Sheepskin Effects in the Return to Education", Review of Economics and Statistics, 69 No 1, (1987): 175-177.

Lang, Kevin and David Kropp, "Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws",Quarterly Journal of Economics, 101 No. 3 (August 1986): 609-624.

Manski, Charles F., "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon", Economics of Education Review, 8 No. 4, (1989): 305-312.

Mincer, Jacob, "Education and Unemployment", National Center on Education and Employment, Teachers College Columbia University, (1988).

Mincer, Jacob, Schooling, Experience, and Earnings, NBER: New York (1974).

Mincer, Jacob, "On-the-Job Training: Costs, Returns, and Some Implications", Journal of Political Economy (1962)

Mincer, Jacob, "Investment in Human Capital and Personal Income Distribution", Journal of Political Economy August 1958.

Miller, Robert, "Job Matching and Occupational Choice", Journal of Political Economy, 92 (December 1984): 1086-1120.

Murphy, Kevin and Finis Welch, "The Structure of Wages", unpublished paper, Unicon Corporation, (September 1988).

Paglin, Morton and Anthony M. Rufolo, "Heterogenous Human Capital, Occupational Choice, and Male-Female Earnings Differences", Journal of Labor Economics 8 No 1, Part 1, (January 1990): 123-144.

Polachek, Solomon, "Occupational Self Selection: A human capital approach to Sex Differences in Occupational Structure", Review of Economics and Statistics 58 pp. 60-69.

Rosen, Sherwin (1977), "Human Capital: A Survey of Empirical Research." In R. Ehrenberg ed., Research in Labor Economics, Greenwich, CT: JAI Press.

Shaw, Kathryn L., "Occupational Change, Employer Change, and the Transferability of Skills" Southern Economic Journal (1987): 702-719.

Sicherman, Nachum and Oded Galor: "A Theory of Career Mobility", Journal of Political Economy, 98 No. 1 (February 1990): 169-192.

Weisbrod, Burton A, "Education and Investment in Human Capital", Journal of Political Economy, Supplement, LXX (October 1962):106-123.

Willis, Robert J. and Sherwin Rosen, "Education and Self Selection", Journal of Political Economy, 87(2) (1979):S7-S36

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in O.Ashenfelter and R. Layard (editors) Handbook of Labor Economics (1987).

Appendix 1: Proof of Proposition 1

I first examine the properties of $V_2(g_2, j)$ and then use the assumption that F_{hm} and F_h are stochastically increasing in K_0 and A to establish properties of $E(V_2(g_2, j)|K_0, A)$. Proposition 1 follows almost immediately.

Lemma A1.1: $V2(g_2, 2)$ is nondecreasing in g_{2h}, g_{2m}, strictly increasing in Y_1, and nondecreasing in Y_{2h}.

Lemma A1.2: $V2(g_2, 1)$ is nondecreasing in g_{2h}, g_{2m}, strictly increasing in Y_1, and nondecreasing in Y_{2h} and Y_{2m}.

Proof: Recall that $g_2=(g_{2h}, g_{2m})$. Both lemmas are obvious upon inspection of equations (6) and (7), since the max function is nondecreasing in all arguments and Y_{2h} and Y_{2m} are assumed greater than $Y_1/(1 + R)$.

Lemma A1.3: $EV_2(K_1,A,2 \mid K_0, A, c_1)/Y_0$, c = m or h, is strictly increasing in K_0 and A, Y_{2h}/Y_0, Y_{2m}/Y_0 and Y_1/Y_0. It is strictly decreasing in R.

Proof: The first argument of the Max( ) function in on the left side of (7) is strictly increasing in Y_{2h} and g2h (given $Y_{2h} > Y_1/(1 + R)$) and strictly decreasing in R. The second argument is strictly increasing in g_{2m}, Y_{2m}, and Y_1 and strictly decreasing in R. The third argument is strictly increasing in Y_1. Consequently, the distribution of at least one of the three arguments is strictly increasing in Y_{2h}, Y_{2m}, and Y_1 and strictly decreasing in R. Furthermore, the distribution F_{hm} of (g_{2h},g_{2m}) is assumed to be stochastically increasing in K_0 and A, which implies that the distribution of the arguments that are strictly increasing in g_{2h} and/or g_{2m} are stochastically increasing in K_0 and A. These facts and the fact that positive shifts in the distribution of one or more elements of a set of random variables increases the expected value of the maximum of the set establishes the results in the lemma.

The same line of reasoning establishes

Lemma A1.4: $EV_2(K_1, A, 1 \mid K_0, A, c_1)/Y_0$, $c = m$ or h is strictly increasing in K_0, A, Y_{2h}/Y_0, and Y_1/Y_0. It is strictly decreasing in R.

Proposition 1 follows from Lemma A1.3 and A1.4 once one notes that the value of starting college is the maximum of the value of starting college in m or in h and recalls that

$$V_1(K_0, A, \theta, m_1) = \sum_{j=0}^{2} \theta_j \, E(V_2(g_2, j) \mid K_0, A, m_1) ,$$

and

$$V_1(K_0, A, \theta, h_1) = \sum_{j=0}^{2} \theta_j \, E(V_2(g_2, j) \mid K_0, A, h_1) .$$

Appendix 2: The Post Secondary Education Outcome Probabilities

All type 0 individuals drop out of college. Type 1 individuals drop out of college if

(A2.1) $$\frac{g_{2h}(1+r'_{2h})}{(1+R)^2} + \frac{(1-g_{2h})}{(1+R)^2} < \frac{1}{(1+R)}$$

or if $g_{2h} < (R/r'_{2h})$, where $(1+r'_{2c}) = (1+r_{2c})/(1+r_1), c = h$ or m.

Type 2 individuals would choose to drop out rather than spend a second year as a humanities major if

$$g_{2h} < (R/r'_{2h})$$

Using an equation analogous to (A2.1) one concludes that type 2 individuals would prefer to drop out rather than major in math if $g_{2m} < (R/r'_{2m})$. Consequently, they drop out of school if both $g_{2m} < (R/r'_{2m})$ and $g_{2h} < (R/r'_{2h})$. Thus,

(A2.2) $$P_{1c1.} = \theta_0 + (1-\theta_0)(1-\theta'_2)F_m\left(\frac{R}{r'_{2h}} | K_0, A\right) + (1-\theta_0)(\theta'_2)F_{hm}\left(\frac{R}{r'_{2h}}, \frac{R}{r'_{2m}} | K_0, A\right)$$

where F_h is the marginal CDF of g_{2h} and F_{hm} is the joint CDF of g_{2h}, g_{2m}, and $\theta'_2 = \theta_2/(\theta_1+\theta_2)$. I have suppressed the fact that both F_m and F_{hm} are conditional on the first period field c.

It is obvious from inspection of (A2.2) that $P_{1c1.}$ is increasing in θ_0. Since $F_m > F_{hm}$, $P_{1c1.}$ is also decreasing in θ'_2. The other results in

Proposition 3.a follow from the fact that $P_{1c1.}$ is increasing in the value of F_m and F_{hm} and the fact that the monotone likelihood ratio property implies that these functions are decreasing in K_0 and A.

The probability $P_{1c2.}$ is the probability that a person is type 1 or type 2 has a sufficiently high graduation probability to attempt to complete college but is unsuccessful in doing so. The equation is

$$
\begin{aligned}
P_{1c2.} = {} & (1-\theta_0)(1-\theta_2')\int_{R/r_{2h}'}^{1}(1-g_{2h})\,dF_h(g_{2h}|K_0,A) \\
& + (1-\theta_0)\theta_2'\left\{\int_{R/r_{2h}'}^{1}\int_{g_{2h}r_{2h}'/r_{2m}'}^{1}(1-g_{2m})\,dF_{hm}(g_{2h},g_{2m}|K_0,A)\right. \\
& \left. + \int_{0}^{R/r_{2h}'}\int_{R/r_{2m}'}^{1}(1-g_{2m})\,dF_{hm}(g_{2h}g_{2m}|K_0,A)\right\} \\
& + (1-\theta_0)\theta_2'\int_{R/r_{2h}'}^{1}\int_{0}^{g_{2h}r_{2h}'/r_{2m}'}(1-g_{2h})\,dF_{hm}(g_{2h},g_{2m}|K_0,A)
\end{aligned}
\tag{A2.3}
$$

The first term is the probability that the person is type 1, has a graduation probability above the critical value R/r_{2h}', but then fails to meet the requirements in the humanities. The next term is the probability that the individual is type 2, attempts to graduate in m, and is unsucessful. The last term is probability that the person is type 2, attempts to graduate in humanities, and fails to do so. The ranges of integration reflect the fact that type 2 individuals are best off in an expected value sense from choosing m if

$$g_{2m} > R/r_{2m}' \ \underline{\text{and}}\ g_{2m} > g_{2h}r_{2h}'/r_{2m}'.$$

They are best off choosing h if

$$g_{2h} > R/r'_{2h} \underline{\text{ and }} g_{2m} < g_{2h}r'_{2h}/r'_{2m}.$$

It is obvious that $P_{1c2.}$ is decreasing in θ_0. An increase in r'_{2m} unambiguously increases the term on the 3rd line of (A2.3). If $r'_{2m} > r'_{2h}$, it is easy to show that the derivative with respect to r'_{2m} of the sum of the terms on the 2nd and 4th lines of (A2.3) is positive. (The increase in r'_{2m} induces some students to move from humanities to math even though $g_{2h} > g_{2m}$. This raises $P_{1c2.}$) This proves that $P_{1c2.}$ is increasing in r'_{2m} if $r'_{2m} > r'_{2h}$. Using the same methods it is easy to show that $P_{1c2.}$ is increasing in both r'_{2m} and r'_{2h} if $r'_{2m} = r'_{2h}$. An increase in R works in the opposite direction, reducing the fraction of students who attend college in the second year, $P_{1c2.}$.

Without further assumptions about F_h and F_{hm}, the effects of increases in K_0, A, and θ'_2 are ambigous for reasons discussed in the text.

I now turn to Proposition 3.3. The equation for $P_{1c2h} + P_{1c2m}$ is 1 minus the expressions for $P_{1c1.}$ and $P_{1c2.}$ in (A2.1) and (A2.2). For completeness, I present the equations for P_{1c2h} and P_{1c2m} below:

(A2.4)

$$P_{1c2h} = (1-\theta_0)1-\theta'_2 \int_{R/r'_{2h}}^{1} g_{2h} dF_h(g_{2h}|K_0, A)$$

$$+ (1-\theta_0)\theta'_2 \int_{R/r_{2h}}^{1} \int_{0}^{g_{2h}r'_{2h}/r'_{2m}} g_{2h} dF_{hm}(g_{2h}, g_{2m}|K_0, A)$$

The first term of (A2.4) is θ_1 times $E(g_{2h}|g_{2h}>R/r'_{2h})\cdot Prob(g_{2h}>R/r'_{2h})$. The second term is θ_2 times $E(g_{2h}|g_{2m}<g_{2h}r'_{2h}r2m', g_{2h}>R/r'_{2h})\cdot Prob(g_{2m}<g_{2h}r'_{2h}r_{2m}, g_{2h}>R/r'_{2h})$. The equation for P_{1c2m} is:

$$\text{(A2.5)}\quad \begin{aligned} P_{1c2m} = {} & (1-\Theta_0)\Theta'_2 \int_0^{r/r'_{2h}} \int_{R/r'_{2m}}^{1} g_{2m} dF_{hm}(g_{2h}, g_{2m}|K_0, A) \\ & + (1-\Theta_0)\Theta'_2 \int_{R/r'_{2h}}^{1} \int_{g_{2h}r'_{2h}/r'_{2m}}^{1} g_{2m} dF_{hm}(g_{2h}, g_{2m}|K_0, A) \quad . \end{aligned}$$

$P_{1c2h} + P_{1c2m}$ is decreasing in Θ_0 because all type 0 persons drop out after the first year. An increase in R reduces the fraction of students who attend college in the second year, lowering both P_{1c2h} and P_{1c2m}.

To prove that $P_{1c2h} + P_{1c2m}$ is increasing in A and K_0 when $r'_{2h} = r'_{2m}$, note that in this case one may combine (A2.4) and (A2.5) and rewrite them as

$$\text{(A2.6)}\quad \begin{aligned} P_{1c2h} + P_{1c2m} = {} & (1-\Theta_0)(1-\Theta'_2) \int_{R/r'_{2m}}^{1} g_{2h} dF_h(g_{2h}|K_0, A) \\ & + (1-\Theta_0)\Theta'_2 \int_0^1 \int_0^1 \max\left(0, g_{2h} - \frac{R}{r'_{2m}}, g_{2m} - \frac{R}{r'_{2m}}\right) dF_{hm}(g_{2h}, g_{2m}|K_0, A) \\ & + (1-\Theta_0)\Theta'_2 \left[1 - F_{hm}\left(\frac{R}{r'_{2m}}, \frac{R}{r'_{2m}}|K_0, A\right)\right] \frac{R}{r'_{2m}} \end{aligned}$$

The assumption that CDFs F_h and F_{hm} have the monotone likelihood ratio property in A and K_0 implies almost immediately that the first term is increasing in A and K_0. The fact that $\max(0, g_{2h} - R/r'_{2h}, g_{2m} - R/r'_{2m})$ is nondecreasing in g_{2h} and g_{2m} and strictly increasing over part of the range between 0 and 1 and the assumption about F_{hm} implies that the second term is strictly increasing in A and K_0. The assumption about F_{hm} implies that $1 - F_{hm}(R/r'_{2m}, R/r'_{2m}|K_0, A)$ is strictly increasing in K_0 and A. Consequently, the third term in (A2.6) is also strictly increasing in K_0 and A.

The positive effect of θ'_2 on the $P_{1c2h} + P_{1c2m}$ when $r'_{2h} = r'_{2m}$ follows from the fact that the weight on $(1-\theta'_2)$ is $(1-\theta_0)$ times the product of the expectation of g_{2h} given $g_{2h} > r/r'_{2m}$ and the probability that $g_{2h} > r/r'_{2m}$. This weight is strictly less than the weight on θ'_2, which is $(1 - \theta_0)$ times the product of the expectation of $\max(g_{2h}, g_{2m})$ given that $\max(g_{2h}, g_{2m}) > R/r'_{2m}$ and $\text{Prob}(\max(g_{2h}, g_{2m}) > R/r'_{2m})$.

Appendix 3: The Expected Ex Post Return to College When Returns are Uncertain.

In this appendix I show that variables that alter the ex ante probability of completing college, such as ability or high school preparation, will be related ex post to the earnings conditional on completing college even if the variables are not related to the unconditional distribution of the return to college. I work with the model in Section 2 but impose the simplifying assumption that there is only one field of study in college with return R_2. Let $F(g_2|K_0, A)$ be the cdf of the probability g_2 of completing graduation requirements conditional on K_0 and A. I assume that this CDF has the monotone likelihood ratio property in K_0 and A. Since there is only one field of college I need only distinguish between type 0 and type 1 individuals. Type 0 individuals dislike college and will drop out of college after the first year regardless of g and R_2. Type 1 individuals have no nonpecuniary preference between working and college.

I drop the assumption that the returns to college are nonstochastic. Let R_2 denote the return to a college degree. I assume that R_2 is equal to

(A3.1) $$R_2 = r_2(K_0, A) + \phi,$$

where ϕ has a continuous distribution over the range $(0,\bar{\phi})$ and is independent of K_0, A, θ, g_2, and whether the individual will complete graduation requirements conditional on g_2. At the end of the first year of college, the individual learns ϕ. For simplicity, I analyze the case in which the nonstochastic portion $r_2(K_0, A)$ of R_2 is a constant which I set to 0. I also set Y_1 to $Y_0/(1 + R)$. I discuss the case in which $r_2(K_0,A)$ depends on K_0 and A at the end of the section.

Let $E(\phi|K_0, A, \text{graduation})$ be the expectation of ϕ conditional on graduation from college, K_0, and A. Under the assumptions made earlier, the following proposition holds.

Proposition A3: $E(\phi|K_0, A, \text{graduation})$ is decreasing in K_0 and A.

Note that by assumption K_0 and A are independent of ϕ. Furthermore, both are determined prior to starting college and prior to the realization of ϕ. The proposition holds because K_0 and A increase the probability of completing graduation requirements. This lowers the critical value of ϕ required to justify staying in college a second year of college and taking the risk of failing to meet the requirements.

To formally prove Proposition A3, note that the value function at the end of the first period of college is

$$\text{(A3.2)} \qquad V_2(g_2,\phi,1) = \frac{Y_0}{1+R}\max\left\{\left[g_2\frac{(1+\phi)}{(1+R)} + (1-g_2)\frac{1}{(1+R)}\right],1\right\}$$

for individuals who turn out to be type 1. (The value function for type 0 individuals is irrelevant, since these individuals always drop out).

A type 1 individual prefers to remain in school if $\phi > R/g_2$.

Given the assumption that with $\bar{\phi} > R$ so that education is profitable if the person knows she is type 1, will receive $\bar{\phi}$, and is certain to graduate ($g_2 = 1$), then

(A3.3) $\quad E(\phi|g_2, \text{graduate}) = E(\phi|\phi > R/g_2)$,

which is strictly decreasing in g_2. It remains to evaluate the expectation of ϕ conditional on K_0, A, and college graduation, which is

(A3.4) $\quad E(E(\phi|\phi > R/g_2)|K_0, A, \text{graduate}) = E(\phi|K_0, A, \text{graduate})$.

The above expectation is

(A3.5) $$E(\phi|K_0, A, \text{graduate}) = \int_0^1 E(\phi|\phi > R/g_2)\, dF^*(g_2|K_0, A)$$

where dF^* is the product of the probability density of college graduation conditional on g_2 (and type $j = 1$) and the density of g_2 given K_0 and A. dF^* is

(A3.6) $$dF^*(g_2|K_0,A) = Prob(\phi > R/g_2) g_2\, dF(g_2|K_0,A)/\Omega(K_0,A)$$

where $\Omega(K_0,A) = \int_0^1 Prob(\phi > R/g_2)\, g_2 dF(g_2|K_0,A)$.

Since from (A3.6)

$dF^*(g_2|K_0',A')/dF^*(g_2/K_0'',A'') = [dF(g_2/K_0',A')/dF(g_2/K_0'',A'')][\Omega(K_0,'A_0')/\Omega(K_0'',A'')$,

the assumption that dF has the monotone likelihood ratio property implies that dF^* has the monotone likelihood ratio property. Since $E(\phi|\phi > R/g_2)$ is decreasing in R/g_2, it follows that $E(\phi|K_0, A, \text{graduate})$ is decreasing in K_0 and A, which completes the proof.

The proposition does not examine the effects of nonpecuniary preferences. In the above model variables that alter θ_0 and θ_1 do not alter the expectation of ϕ conditional on K_0, A, and graduation. θ_0 and θ_1 may alter the expectation of ϕ conditional only on graduation by altering the distribution of K_0 and A among those who choose to start college. If one extends the model to include a third preference type who "likes" (receives a nonpecuniary benefit from attending) college, then one may show that the expected value of ϕ is negatively related to factors that increase the ex ante probability that the individual will like college even after one conditions on K_0 and A.

If $r_2(K_0, A)$ is a positive function of K_0 and A, then the proposition remains true, since the increase in r_2 lowers the critical value of ϕ at which the individual decides to drop out. If elements of K_0 and A lower $r_2(K_0, A)$, then the overall effect of increases in these elements on $E(\phi | K_0, A,$ graduation) is ambiguous.

Table 1

THE INTERNAL RATE OF RETURN TO STARTING COLLEGE (ρ)

Sample Used to Estimate and to Evaluate the Wage and Education Equations:[a]	(1) Men and Women	(2) Men and Women	(3) Men	(4) Men	(5) Women	(6) Women
1.A Means for All Variables[a]						
EXP*s Slope (α)						
.000[b]	5.1		2.8		7.4	
.0011	6.2		4.1		8.4	
.005	9.0		7.3		10.8	
1.B Ability Level[c]						
	HIGH	LOW	HIGH	LOW	HIGH	LOW
EXP*s Slope (α)						
.000[b]	6.4	3.6	4.2	0.5	8.8	5.9
.0011	7.4	4.9	5.4	2.1	9.5	7.0
.005	9.9	8.0	8.3	5.7	11.7	9.8
1.C Family Background[d]						
	FAV	UNFAV	FAV	UNFAV	FAV	UNFAV
EXP*s Slope (α)						
.000[b]	5.4	4.8	3.3	1.9	7.3	7.6
.0011	6.4	5.9	4.5	3.3	8.3	8.5
.005	9.2	8.7	7.6	6.6	10.7	10.8
1.D High School Curriculum[e]						
	AC	NONAC	AC	NONAC	AC	NONAC
EXP*s Slope (α)						
.000[b]	5.5	4.8	3.1	2.1	7.7	7.6
.0011	6.6	5.9	4.4	3.5	8.6	8.5
.005	9.3	8.7	7.5	6.8	11.0	10.8
1.E Started College?[f]						
	Yes	NO	YES	NO	YES	NO
EXP*s Slope (α)						
.000[b]	5.6	4.1	3.5	0.9	7.8	6.8
.0011	6.7	5.3	4.7	2.4	8.7	7.8
.005	9.4	8.3	7.8	6.0	11.0	10.3

1.F College Plans[g]

	YES	NO	YES	NO	YES	NO
EXP*s Slope (α)						
.000[b]	5.3	4.7	3.1	1.9	7.3	7.7
.0011	6.4	5.8	4.4	3.3	8.2	8.6
.005	9.2	8.6	7.5	6.6	10.7	10.9

a. The text for the specification of the wage model and the 18 probit models for the education outcomes. The internal rate of return is computed using the estimates of these equations evaluated at the sample means of all variables unless otherwise noted in the title for each panel of the table.

b. The Exp*s slope coefficient α is the coefficient on the product of years of academic education and experience. The wage equation was estimated with this parameter set to .000, .0011 and .005 respectively. The alternative estimates of the wage equation were used to evaluate the rate of return.

c. The wage and education equations are evaluated at the sample means of all variables except those related to ability, which are set to either high or low values. In the case of "High" ("Low") ability SAT_VERBAL, SAT_MATH, GRADES, and -COL_ABIL are set to the mean for the particular sample (pooled, men, or women) plus (minus) one standard deviation. The standard deviations are taken from the full sample, so the difference in the values of the aptitude measures for the high and low aptitude cases is the same regardless of whether the group analyzed is men and women, men, or women.

d. The wage and education equations are evaluated at the sample means of all variables except the family background variables, which are set to either high or low values. In the case of favorable background father's education, mother's education, and the log of family income are set to the mean for the particular sample (pooled, men, or women) plus (minus) one standard deviation. The dummies for whether the individual discussed plans with parents and whether he/she was influenced by parents were set to 1. Dummy variables for whether lack of parental interest interfered with high school education and for whether lack of money interfered with high school education were set to 0. The definition of "unfavorable" is symmetric to the favorable case. The standard deviations are taken from the full sample.

e. The wage and education equations are evaluated at the sample means of all variables except those related to high school curriculum. In the case of AC, semester hours of math, science, and foreign language are the means for the particular sample (pooled, men, or woemn) plus one standard deviation. Industrial arts and commercial were set to the sample means minus one standard deviation. In the case of NONAC, the opposite adjustment was made. The standard deviations of the curriculum variables are taken from the particular sample (pooled, men and women). I do so because the standard deviations of industrial arts and commercial courses differ dramatically between men and women. In the case of AC, the dummy variable for whether the individual was in an academic program was set to one and the dummy for a general high school program was set to 0. (Vocational is the reference group in all equations.) In the case of NONAC, the dummy for academic program was set to 0 and the dummy for general high school program was set to the conditional probability that an individual is in a general program given that he or she is not in the academic program.

f. The wage and education equations are evaluated using the sample means and women alone of those who started college (Yes) and the sample means for those who did not start college (No).

g. The wage and education equations are evaluated at the sample means of all variables except those related to college plans. The education equations were augmented to include a series of measures of educational plans. Individuals are defined to have 4 year college plans ("Yes") if they indicated they planned to complete college and if they indicated that they planned to attend or to transfer to a 4 year college. They are defined not to have 4 year college plans if they indicated that they did not plan to complete college and did not plan to attend or to transfer to a 4 year college.

Table A.1
Log Wage Coefficients and Sample Probabilities for Various Post Secondary Education Outcomes.

Education Outcome	Men and Women		Men		Women	
	Wage Coef.[a] (1)	Prob.[b] (2)	Wage Coef.[a] (3)	Prob.[b] (4)	Wage Coef.[a] (5)	Prob.[b] (6)
College, LT 2	0.0369 (0.0116)	.286	-0.0076 (0.0180)	.264	0.0789 (0.0142)	.308
College, 2+	0.1292 (0.0128)	.272	0.0575 (0.0200)	.287	0.1945 (0.0174)	.258
College Degree						
Business, communications	0.3153 (0.0215)	.067	0.2113 (0.0300)	.077	0.4647 (0.0337)	.039
Engineering	0.5105 (0.0290)	.017	0.4436 (0.0353)	.031	0.5061 (0.1012)	.002
Physical sciences	0.3367 (0.0680)	.009	0.2749 (0.0842)	.014	0.2968 (0.0778)	.004
Humanities	0.1676 (0.0312)	.023	0.0963 (0.0556)	.018	0.2177 (0.0361)	.027
Social sciences, law services	0.1939 (0.0209)	.086	0.1292 (0.0345)	.089	0.2342 (0.0252)	.083
Life Sciences, health	0.3133 (0.0215)	.075	0.1548 (0.0347)	.074	0.4282 (0.0254)	.077
Education, home economics, library science	0.1588 (0.0209)	.070	0.0317 (0.0427)	.038	0.2225 (0.0224)	.101
Math, computer science	0.4475 (0.0480)	.006	0.4204 (0.0739)	.006	0.4550 (0.0648)	.006
Fine Arts	0.1431 (0.0414)	.014	0.0256 (0.0617)	.011	0.2281 (0.0471)	.018
Miscellaneous, Field missing	0.2278 (0.0281)	.031	0.2327 (0.0454)	.029	0.3024 (0.0310)	.033
Advanced						
Math, physical science, engineering, computer science	0.5369 (0.0725)	.0047	0.4234 (0.0786)	.0076	0.8470 (0.0515)	.0018
Business law, communications	0.5138 (0.0541)	.0077	0.4302 (0.0611)	.0110	0.5982 (0.1154)	.0044
Humanities, social	0.3497 (0.0550)	.0068	0.2108 (0.0515)	.0067	0.4267 (0.0861)	.0068
Life sciences, health	0.3916 (0.0711)	.0071	0.2725 (0.1137)	.0073	0.4960 (0.0767)	.0068
Education, home economics, library science	0.3436 (0.0556)	.0117	0.2040 (0.0977)	.0046	0.4368 (0.0624)	.0186
Arts, services, field of degree missing	0.3181 (0.0927)	.0060	0.3194 (0.1618)	.0073	0.2973 (0.0829)	.0047

a) The other variables in the wage equation are described on page 26 and 27. Interactions terms involving the education outcomes are included. The standard errors in parenthesis allow for arbitrary forms of heteroscedasticity and correlation among residuals for a given individual or individuals from the same high school. The R2 for columns 1, 3, and 5 are 0.1958, 0.1175 and 0.2062 respectively.

b) The sample probabilities are conditional upon starting college.

Table A.2

Distribution of College Majors
and Advanced Degrees

Majors College Major	% of Males With This Major	% of Majors Who are Male	% of Females With This Major	% of Who are Female
	Individuals with College or College and Advanced Degrees			
Business communications	22.84	71.19	9.24	28.81
Engineering	7.96	93.60	.54	6.40
Physicial science	3.26	72.73	1.22	27.27
English, foreign language	4.68	39.88	7.06	60.12
Sociology, psychology, law	21.74	50.47	21.34	49.53
Life sciences	17.38	48.04	18.90	51.96
Education	9.06	26.07	25.82	73.93
Math, computer science	1.83	54.00	1.56	46.00
Fine arts	2.50	36.27	4.42	63.73
Missing	9.06	49.26	9.38	50.74
Total	1479		1464	
Advanced Degree	Individuals with Advanced Degrees			
Math, physical science, engineering	17.12	80.65	4.10	19.35
Business, law, communications	24.66	70.59	10.28	29.41
humanities, social sciences	15.06	48.89	15.76	51.11
Life sciences	16.44	51.06	15.76	48.94
Education	10.28	19.23	43.16	80.77
Missing	16.44	60.00	10.96	40.00
Total advanced degree	146		146	

Table A.3
DESCRIPTIVE STATISTICS, MEN AND WOMEN

VARIABLE	LABEL	FULL SAMPLE MEAN	FULL SAMPLE STANDARD DEV.	NO COLLEGE MEAN	NO COLLEGE STANDARD DEV.	SOME COLLEGE MEAN	SOME COLLEGE STANDARD DEV.
BLACK	BLACK, AMER.INDIAN	0.09926	0.29902	0.11419	0.31809	0.11350	0.31729
HISP	MEXICAN,PUERTORICO,LATIN AMERICAN	0.03541	0.18483	0.04292	0.20272	0.04203	0.20073
CSEX	SEX COMPOSITE: MALE = 0, FEMALE = 1	0.52658	0.49931	0.56136	0.49628	0.54755	0.49786
Family Background							
CFAED	FATHERS EDUCATION, COMPOSITE	12.60081	2.50669	11.50632	1.86514	12.37729	2.33533
CMOED	MOTHERS EDUCATION, COMPOSITE	12.32621	2.08474	11.51238	1.64909	12.14293	1.95010
LOGINC	LOG FAMILY INCOME IN 1972	8.974	0.586	8.784	0.569	8.936	0.578
BQ17JMON	MONEY PROBLEMS INTERFERE WITH ED	0.28827	0.45298	0.36048	0.48020	0.31213	0.46348
MWRKE	MOTHER WORKED WHEN IN ELEM.SCHOOL	0.40238	0.49040	0.40368	0.49070	0.44771	0.49738
PLANDSCD	PARENTS DISCUSSED FUTURE PLANS	0.78662	0.40970	0.70638	0.45548	0.76510	0.42404
PLANINFL	PARENTS INFUENCED FUTURE PLANS	0.43867	0.49624	0.37259	0.48356	0.39358	0.48867
BQ17FUIP	UNINTERESTED PARENTS INTERFERE W ED.	0.21395	0.41011	0.30902	0.46215	0.24750	0.43167
High School Program							
BQ7	TIME/WK SPENT ON HOMEWORK	4.41934	3.26278	3.63856	2.88145	4.01146	3.03293
HSACAD	HS-PROGRAM IS ACADEMIC (1=YES,0=NO)	0.45216	0.49773	0.16318	0.36958	0.38045	0.48562
HSGEN	HS-PROGRAM IS GENERAL (1=YES,0=NO)	0.31496	0.46452	0.41001	0.49190	0.37046	0.48305
COLLPROX	SCHOOL PROXIMITY TO COLLEGE	1.79352	0.75499	1.93482	0.78172	1.70833	0.74410
PRFQTHR1	SEM. HRS SCIENCE , PRED.	18.72539	10.08141	14.83891	8.79246	17.48391	9.66468
PRFQTHR2	SEM. HRS FOREIGN LANGUAGES, PRED	10.84675	11.32240	5.21479	8.65498	9.61250	10.38436
PRFQTHR3	SEM. HRS SOCIAL STUDIES, PRED.	26.23072	7.67122	26.00046	7.74759	26.12846	7.52894
PRFQTHR4	SEM. HRS ENGLISH , PRED	29.99487	6.60730	29.34478	6.47051	29.73800	6.57107
PRFQTHR5	SEM. HRS MATHEMATICS, PRED	19.17939	9.99209	14.61655	9.30781	18.20686	9.49172
PRFQTHR6	SEM. HRS INDUSTRIAL ARTS, PRED	5.61403	11.90166	7.54189	14.16043	5.98975	12.40059
PRFQTHR7	SEM. HRS COMMERCIAL STUDIES,PRED	14.05415	16.50500	19.20508	19.85767	15.06001	16.04808
PRFQTHR8	SEM. HRS FINE ARTS, PRED	8.68579	13.17194	7.24458	11.85174	9.35001	13.36491
Aptitude and Achievement Measures							
SATM_HAT	PREDICTED SAT SCORE: MATH	427.42329	100.86270	370.56558	75.29589	409.76767	86.85755
SATV_HAT	PREDICTED SAT SCORE: VERBAL	397.96923	97.42458	344.32919	74.17472	383.36516	83.23649
BQ28	DOES STUD. BELIEVE HE HAS COLL. ABIL.	1.88734	0.98677	2.44744	1.09906	1.86442	0.89131
BQ5	GRADES IN HIGH SCHOOL	80.95753	7.58272	77.89812	7.07144	79.45181	6.84694
Location							
SMLTOWN	STUDENT COMES FROM SMALL HOMETOWN	0.27760	0.44783	0.27820	0.44817	0.27798	0.44812
MEDCITY	STUDENT COMES FROM MEDIUM CITY	0.11866	0.32341	0.12080	0.32594	0.13137	0.33789
MEDSURB	STUDENT COMES FROM SUBURB OF MED CITY	0.07869	0.26927	0.07099	0.25685	0.08197	0.27440
BIGCITY	STUDENT COMES FROM BIG CITY	0.08519	0.27918	0.07292	0.26004	0.09406	0.29199
BIGSURB	STUDENT COMES FROM SUBURB OF BIG CITY	0.08723	0.28218	0.05971	0.23698	0.09090	0.28755
HUGECITY	STUDENT COMES FROM HUGE CITY	0.04773	0.21322	0.03714	0.18915	0.05465	0.22735
HUGESURB	STUDENT COMES FROM SUBURB OF HUGE CITY	0.06782	0.25145	0.03082	0.17285	0.06305	0.24313
NOCENTR	REGION=NORTH-CENTRAL	0.28594	0.45188	0.30599	0.46089	0.28849	0.45318
SOUTH	REGION=SOUTH	0.32621	0.46885	0.35195	0.47764	0.33158	0.47090
WEST	REGION=WEST	0.16873	0.37453	0.13181	0.33833	0.21019	0.40755

Table A.3 Continued

COLLEGE OR ADVANCED

VARIABLE	LABEL	MEAN	STANDARD DEVIATION
BLACK	BLACK, AMER.INDIAN	0.06863	0.25287
HISP	MEXICAN,PUERTORICO,LATIN AMERICAN	0.01664	0.12797
CSEX	SEX COMPOSITE: MALE = 0, FEMALE = 1	0.49745	0.50007
Family Background			
CFAED	FATHERS EDUCATION, COMPOSITE	13.88854	2.86784
CMOED	MOTHERS EDUCATION, COMPOSITE	13.29323	2.20082
LOGINC	LOG FAMILY INCOME IN 1972	9.212	0.533
BQ17JMON	MONEY PROBLEMS INTERFERE WITH ED	0.18484	0.38823
MWRKE	MOTHER WORKED WHEN IN ELEM.SCHOOL	0.36493	0.48149
PLANDSCD	PARENTS DISCUSSED FUTURE PLANS	0.88005	0.32495
PLANINFL	PARENTS INFUENCED FUTURE PLANS	0.53244	0.49903
BQ17FUIP	UNINTERESTED PARENTS INTERFERE EDUC/NO=0	0.09106	0.28774
High School Program			
BQ7	TIME/WK SPENT ON HOMEWORK	5.58137	3.53350
HSACAD	HS-PROGRAM IS ACADEMIC (1=YES,0=NO)	0.79816	0.40143
HSGEN	HS-PROGRAM IS GENERAL (1=YES,0=NO)	0.17023	0.37590
COLLPROX	SCHOOL PROXIMITY TO COLLEGE	1.72349	0.72879
PRFQTHR1	SEM. HRS SCIENCE , PRED.	23.52335	9.80980
PRFQTHR2	SEM. HRS FOREIGN LANGUAGES, PRED	17.58599	11.17244
PRFQTHR3	SEM. HRS SOCIAL STUDIES, PRED.	28.42870	7.62580
PRFQTHR4	SEM. HRS ENGLISH , PRED	30.90932	6.59548
PRFQTHR5	SEM. HRS MATHEMATICS, PRED	24.49334	8.47519
PRFQTHR6	SEM. HRS INDUSTRIAL ARTS, PRED	2.95682	7.56419
PRFQTHR7	SEM. HRS COMMERCIAL STUDIES,PRED	8.52875	10.63317
PRFQTHR8	SEM. HRS FINE ARTS, PRED	9.89006	14.34163
Aptitude and Achievement Measures			
SATM_HAT	PREDICTED SAT SCORE: MATH	502.65626	93.38801
SATV_HAT	PREDICTED SAT SCORE: VERBAL	467.50043	92.86582
BQ5	GRADES IN HIGH SCHOOL	85.61153	6.84981
BQ28	DOES STUD. BELIEVE HE HAS COLL. ABIL.	1.36119	0.58564
Location			
SMLTOWN	STUDENT COMES FROM SMALL HOMETOWN	0.28916	0.45344
MEDCITY	STUDENT COMES FROM MEDIUM CITY	0.10771	0.31007
MEDSURB	STUDENT COMES FROM SUBURB OF MED CITY	0.07849	0.26898
BIGCITY	STUDENT COMES FROM BIG CITY	0.08868	0.28433
BIGSURB	STUDENT COMES FROM SUBURB OF BIG CITY	0.11348	0.31724
HUGECITY	STUDENT COMES FROM HUGE CITY	0.05028	0.21857
HUGESURB	STUDENT COMES FROM SUBURB OF HUGE CITY	0.11213	0.31558
NOCENTR	REGION=NORTH-CENTRAL	0.27760	0.44789
SOUTH	REGION=SOUTH	0.30071	0.45864
WEST	REGION=WEST	0.14848	0.35564

Table A.4: The Probabilities and Log Wage Gains for Education Outcomes
(Column letters correspond to panels A, B, C, D, E, and F of Table 1)

FULL SAMPLE

CATEGORY	A. Means EX ANTE PROB	A. Means LOGWAGE GAIN	B. High Ability EX ANTE PROB	B. High Ability LOGWAGE GAIN	B. Low Ability EX ANTE PROB	B. Low Ability LOGWAGE GAIN	C. Fav. background EX ANTE PROB	C. Fav. background LOGWAGE GAIN	C. Unfav. Background EX ANTE PROB	C. Unfav. Background LOGWAGE GAIN
COLLEGE, LT 2	.3525	.0409	.2433	.0633	.5418	.0184	.2301	.0409	.5217	.0409
COLLEGE, 2+	.3828	.1307	.3302	.1531	.3531	.1082	.3685	.1307	.3542	.1307
COLLEGE TECH	.0020	.5221	.0028	.4824	.0003	.6252	.0028	.5241	.0010	.5123
COLLEGE NONTECH	.2568	.2242	.4138	.2710	.1044	.1883	.3900	.2285	.1176	.2123
ADVANCE TECH	.0000	.6287	.0000	.5581	.0000	.7844	.0000	.6287	.0000	.6287
ADVANCE NONTECH	.0058	.3628	.0087	.3567	.0003	.4300	.0085	.3754	.0054	.3535
SOME COLLEGE	.7351	.0876	.5735	.1150	.8948	.0538	.5086	.0962	.8758	.0772
COLLEGE	.2588	.2265	.4168	.2725	.1048	.1898	.3928	.2307	.1186	.2148
ADVANCED	.0058	.3628	.0087	.3567	.0003	.4300	.0085	.3754	.0054	.3535

CATEGORY	D. Acad Cur. EX ANTE PROB	D. Acad Cur. LOGWAGE GAIN	D. Nonac Cur. EX ANTE PROB	D. Nonac Cur. LOGWAGE GAIN	E. Start college EX ANTE PROB	E. Start college LOGWAGE GAIN	E. No College EX ANTE PROB	E. No College LOGWAGE GAIN	F. College Plans EX ANTE PROB	F. College Plans LOGWAGE GAIN	F. No-Coll Plans EX ANTE PROB	F. No-Coll Plans LOGWAGE ETURN
COLLEGE, LT 2	.1875	.0408	.5272	.0408	.2434	.0476	.5421	.0285	.2334	.0409	.5688	.0409
COLLEGE, 2+	.3476	.1307	.3800	.1307	.3616	.1374	.3579	.1183	.3198	.1307	.3735	.1307
COLLEGE TECH	.0114	.4668	.0003	.5279	.0046	.4911	.0003	.5751	.0034	.5242	.0004	.5270
COLLEGE NONTECH	.4277	.2276	.0810	.2290	.3739	.2374	.0993	.1896	.4334	.2210	.0570	.2274
ADVANCE TECH	.0000	.6287	.0000	.6287	.0000	.5968	.0000	.7107	.0000	.6287	.0000	.6287
ADVANCE NONTECH	.0158	.3958	.0015	.3652	.0165	.3686	.0004	.3978	.0100	.3676	.0004	.4334
SOME COLLEGE	.5452	.0981	.9072	.0785	.8050	.1013	.9000	.0642	.5532	.0928	.9422	.0765
COLLEGE	.4390	.2338	.0913	.2208	.3785	.2405	.0996	.2007	.4368	.2233	.0574	.2206
ADVANCED	.0158	.3961	.0015	.3652	.0165	.3686	.0004	.3978	.0100	.3676	.0004	.4334

MALES

CATEGORY	A. Means EX ANTE PROB	A. Means LOGWAGE GAIN	B. High Ability EX ANTE PROB	B. High Ability LOGWAGE GAIN	B. Low Ability EX ANTE PROB	B. Low Ability LOGWAGE GAIN	C. Fav. Background EX ANTE PROB	C. Fav. Background LOGWAGE GAIN	C. Unfav. Background EX ANTE PROB	C. Unfav. Background LOGWAGE GAIN
COLLEGE, LT 2	.2858	-.0061	.1881	.0175	.4858	-.0287	.1824	-.0061	.4887	-.0061
COLLEGE, 2+	.4117	.0565	.3313	.0801	.4138	.0328	.3881	.0565	.3907	.0565
COLLEGE TECH	.0126	.4732	.0140	.4849	.0024	.5376	.0190	.4739	.0057	.4681
COLLEGE NONTECH	.2776	.1483	.4642	.1902	.0878	.0895	.4045	.1544	.1147	.1379
ADVANCE TECH	.0000	.4844	.0000	.4494	.0000	.5835	.0000	.4843	.0000	.4843
ADVANCE NONTECH	.0023	.2712	.0023	.3277	.0002	.0266	.0059	.2261	.0002	.3179
SOME COLLEGE	.7075	.0303	.5185	.0575	.9097	-.0013	.5705	.0364	.8784	.0217
COLLEGE	.2902	.1625	.4782	.1982	.0902	.1013	.4236	.1688	.1204	.1535
ADVANCED	.0023	.2712	.0023	.3277	.0002	.0266	.0059	.2261	.0002	.3179

CATEGORY	D. Acad. Cur. EX ANTE PROB	D. Acad. Cur. LOGWAGE GAIN	D. Nonacad. Cur. EX ANTE PROB	D. Nonacad. Cur. LOGWAGE GAIN	E. Start College EX ANTE PROB	E. Start College LOGWAGE GAIN	E. No College EX ANTE PROB	E. No College LOGWAGE GAIN	F. College Plans EX ANTE PROB	F. College Plans LOGWAGE GAIN	F. No-Coll Plans EX ANTE PROB	F. No-Coll Plans LOGWAGE GAIN
COLLEGE, LT 2	.1743	-.0061	.4520	-.0061	.1942	.0014	.5163	-.0215	.2218	-.0061	.4903	-.0061
COLLEGE, 2+	.3787	.0565	.4365	.0565	.3723	.0640	.4000	.0410	.3464	.0565	.4201	.0565
COLLEGE TECH	.0284	.4410	.0048	.4746	.0193	.4590	.0026	.5158	.0188	.4737	.0042	.4746
COLLEGE NONTECH	.4085	.1360	.1061	.1683	.4049	.1591	.0811	.1153	.4092	.1445	.0854	.1559
ADVANCE TECH	.0002	.4843	.0000	.4843	.0000	.4680	.0000	.5402	.0000	.4843	.0000	.4843
ADVANCE NONTECH	.0088	.2120	.0007	.3269	.0093	.2569	.0000	.3111	.0038	.2556	.0000	.4348
SOME COLLEGE	.5530	.0367	.8864	.0246	.5665	.0425	.8163	.0057	.5682	.0320	.9105	.0227
COLLEGE	.4378	.1565	.1109	.1814	.4242	.1728	.0837	.1276	.4280	.1590	.0895	.1707
ADVANCED	.0091	.2165	.0007	.3269	.0093	.2570	.0000	.3111	.0038	.2556	.0000	.4348

Table A4, continued

FEMALES

CATEGORY	A. Means EX ANTE PROB	A. Means LOGWAGE GAIN	B. High Ability EX ANTE PROB	B. High Ability LOGWAGE GAIN	B. Low Ability EX ANTE PROB	B. Low Ability LOGWAGE GAIN	C. Fav. Background EX ANTE PROB	C. Fav. Background LOGWAGE GAIN	C. Unfav. Background EX ANTE PROB	C. Unfav. Background LOGWAGE GAIN
COLLEGE, LT 2	.4060	.0858	.3004	.1162	.5867	.0555	.2771	.0858	.5658	.0858
COLLEGE, 2+	.3562	.1986	.3374	.2290	.2977	.1682	.3470	.1986	.3199	.1986
COLLEGE TECH	.0000	.3781	.0000	.5003	.0000	.2781	.0000	.3781	.0000	.3781
COLLEGE NONTECH	.2366	.2797	.3550	.3311	.1155	.2363	.3729	.2804	.1134	.2684
ADVANCE TECH	.0000	.	.0000	.	.0000	.	.0000	.	.0000	.
ADVANCE NONTECH	.0013	.4907	.0073	.4404	.0001	.5438	.0030	.4979	.0009	.4813
SOME COLLEGE	.7622	.1385	.8377	.1758	.8844	.0934	.6241	.1485	.8857	.1266
COLLEGE	.2368	.2797	.3550	.3311	.1155	.2363	.3729	.2804	.1134	.2684
ADVANCED	.0013	.4907	.0073	.4404	.0001	.5438	.0030	.4979	.0009	.4813

CATEGORY	D. Acad. Cur. EX ANTE PROB	D. Acad. Cur. LOGWAGE GAIN	D. Non-Acad Cur. EX ANTE PROB	D. Non-Acad Cur. LOGWAGE GAIN	E. Start College EX ANTE PROB	E. Start College LOGWAGE GAIN	E. No College EX ANTE PROB	E. No College LOGWAGE GAIN	F. College Plans EX ANTE PROB	F. College Plans LOGWAGE GAIN	F. No Coll Plans EX ANTE PROB	F. No Coll Plans LOGWAGE GAIN
COLLEGE, LT 2	.2153	.0858	.5783	.0858	.2922	.0947	.5734	.0712	.2446	.0858	.6174	.0858
COLLEGE, 2+	.3059	.1986	.3322	.1986	.3507	.2075	.3250	.1839	.2994	.1986	.3361	.1986
COLLEGE TECH	.0000	.3781	.0000	.3781	.0000	.4141	.0000	.3236	.0000	.3781	.0000	.3781
COLLEGE NONTECH	.3980	.2966	.0892	.2749	.3540	.2971	.1013	.2527	.4529	.2776	.0464	.2834
ADVANCE TECH	.0000	.	.0000	.	.0000	.	.0000	.	.0000	.	.0000	.
ADVANCE NONTECH	.0808	.4923	.0003	.4915	.0031	.4710	.0003	.5177	.0031	.4909	.0000	.4957
SOME COLLEGE	.5212	.1520	.9106	.1270	.6429	.1562	.8984	.1120	.5440	.1479	.9535	.1256
COLLEGE	.3980	.2966	.0892	.2749	.3540	.2971	.1013	.2527	.4529	.2778	.0464	.2834
ADVANCED	.0808	.4923	.0003	.4915	.0031	.4710	.0003	.5177	.0031	.4909	.0000	.4957

a) The ex ante probabilities are the probabilities of the various education outcomes. The composite categories COLLEGE NONTECH, COLLEGE TECH, ADVANCE TECH, and ADVANCED NONTECH are defined in the text. All 10 college degree categories and all 5 advanced categories are combined in COLLEGE and ADVANCED. The log wage gain for a particular education category is the weighted averages of the log wage differentials (relative to a high school graduate) for the various degrees in the category using the ex ante probabilities as weights. The values of the individual characteristics used to evaluate the education probabilities and log wage differentials are described in the footnotes to Table 1. See Table 1, fn. a for the A columns, fn. c for the B columns, fn. d for the C columns, footnote e for the D columns, fn. f for the E columns, and fn. g for the F columns.

Table A.5

GENDER DIFFERENCES IN THE INTERNAL RATE OF RETURN TO STARTING COLLEGE (ρ)

		1a	2a	3a	4a	5a	6a	7a	8a	9a	1b	2b	3b	4b	5b	6b
Wage Equation Sample		MF	MF	MF	MF	MF	MF	MF	MF	MF	M	M	M	M	M	M
Education Equations Sample (Only Persons who Started College)		MF	MF	MF	M	M	M	F	F	F	MF	MF	MF	M	M	M
Characteristics Used to Evaluate Education and Wage Equations:		MF	M	F	MF	M	F	MF	M	F	MF	M	F	MF	M	F
EXP*s Slope	-.000[b]	5.1	5.2	5.2	5.6	5.3	5.9	4.8	4.5	5.0	2.4	2.6	2.3	3.0	2.8	3.2
	-.0011	6.2	6.3	6.3	6.7	6.4	6.9	5.9	5.7	6.2	3.7	4.0	3.6	4.3	4.1	4.5
	-.005	9.0	9.1	9.1	9.4	9.2	9.6	8.8	8.6	9.0	7.0	7.2	6.9	7.4	7.3	7.6

		7b	8b	9b	1c	2c	3c	4c	5c	6c	7c	8c	9c
Wage Equation Sample		M	M	M	F	F	F	F	F	F	F	F	F
Education Equations Sample (Only Persons who Started College)		F	F	F	MF	MF	MF	M	M	M	F	F	F
Characteristics Used to Evaluate Eduation and Wage Equations:		MF	M	F	MF	M	F	MF	M	F	MF	M	F
EXP*s Slope	-.000[b]	1.9	1.7	2.1	8.1	8.6	7.6	8.6	8.8	8.4	7.7	8.1	7.4
	-.0011	3.3	3.1	3.4	8.9	9.4	8.5	9.4	9.6	9.2	8.6	8.9	8.4
	-.005	6.6	6.5	6.8	11.2	11.6	10.9	11.6	11.7	11.5	11.0	11.2	10.8

a) MF : Pooled sample of men and women. M: Sample of men. F: Sample of Women. The education outcomes probabilities used in evaluating the internal rate to education are evaluated at the mean of the explanatory variables for the indicated sample. Results using the mean of the probabilities evaluated over the distribution of the explanatory variables are discussed in the text. The wage equations are evaluated at the sample means. See the text for a description of the education and wage equations.

b. The EXP*s slope is the coefficient α on the product of years of academic education and experience. The wage equation was estimated with this parameter set to .000, .0011, and .005 respectively. The alternative estimates of the wage equation were used to evaluate the rate of return.