Volume Title: Studies of Supply and Demand in Higher Education

Volume Author/Editor: Charles T. Clotfelter and Michael Rothschild, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-11054-0

Volume URL: http://www.nber.org/books/clot93-1

Conference Date: May 17-19, 1991

Publication Date: January 1993

Chapter Title: Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?

Chapter Author: Charles F. Manski

Chapter URL: http://www.nber.org/chapters/c6097

Chapter pages in book: (p. 43 - 60)

2 Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?

Charles F. Manski

2.1 Basic Ideas

Economists analyzing schooling decisions assume that youth, having compared the expected outcomes from schooling and other activities, choose the best feasible option. Viewing education as an investment in human capital, we use the term *returns to schooling* to refer to the outcomes from schooling relative to nonschooling.

Given the centrality of the expected returns to schooling in economic thinking on educational behavior, it might be anticipated that economists would make substantial efforts to learn how youth form their expectations. But the profession has traditionally been skeptical of subjective data; so much so that we have generally been unwilling to collect data on expectations. Instead, the norm has been to make assumptions about expectations formation.

2.1.1 Prevailing Expectations Assumptions

Economic studies of schooling behavior have universally assumed that expectations formation is homogeneous; all youth condition their beliefs on the same variables and process their information in the same way. On the other hand, the hypothesized conditioning variables and information processing rule have varied considerably across studies.

In his analysis of the major field decisions of male college students, Freeman (1971) assumed that these youth condition their expectations on their sex

and on their common knowledge of the incomes realized by earlier cohorts. He assumed that expectations formation is myopic. Each youth believes that by selecting a given college major, he will obtain the mean income realized by the members of a specified earlier cohort who made that choice.[1]

Willis and Rosen (1979), in their study of college enrollment, took the personal conditioning variables to be sex, armed forces status, and ability. They assumed that youth have common knowledge of the actual process generating life-cycle incomes conditional on these personal variables and on schooling. They hypothesized that expectations are rational, each youth applying knowledge of the true income-generating process to forecast future personal income should he or she enroll or not enroll in college.

In the Manski and Wise (1983, chap. 6) analysis of college choice, youth condition their expectations for the utility of enrolling in a given college on their own SAT score and on the average SAT score of students enrolled at the college. They do not necessarily know either the outcomes realized by earlier cohorts or the actual process generating outcomes. Rather, they believe the returns from enrolling to be a function of the difference between their own SAT score and the average at the college.

The three studies just cited are noteworthy because they make explicit assumptions about expectations formation. In most economic analyses of schooling behavior, the expectations assumptions are implicit in the specification of the decision model. The recent literature shows little concern with expectations formation. The prevailing sentiment seems to be complacency. Either researchers are confident that their expectations assumptions are correct, or they believe that misspecifying expectations is innocuous.

2.1.2   Two Identification Problems

In fact, there is no evidence that prevailing expectations assumptions are correct nor reason to think that misspecifying expectations is innocuous. To the contrary, rudimentary treatment of expectations has placed the economics of education at an impasse, caught in a pair of basic identification problems that plague attempts to understand schooling behavior and to measure educational productivity.

The first problem is that, not knowing how youth perceive the returns to schooling, one cannot infer their decision processes from their schooling choices. The point can easily be made with a few symbols. The standard economic model assumes that a youth's schooling choice c is a function $f(.)$ of his or her expected returns to schooling r; that is, $c = f(r)$. Suppose that one wishes to learn the decision rule $f(.)$ mapping expectations into choices. If one observes the choices and expectations of a sample of youth, then one can infer

---

1. In the final chapters of his book, Freeman reported findings from a one-time survey of college students regarding their income expectations in various occupations. But his analysis of these data sheds no light on the realism of the myopic expectations assumption made earlier on.

the decision rule. But if one observes only the choices of these youth, then clearly one cannot infer $f(.)$. The most that one can do is infer the decision rule conditional on maintained assumptions on expectations.

The second problem is that, not knowing youth's decision processes, one cannot infer the objective returns to schooling from data on realized outcomes. As is well known, any attempt to learn the objective returns to schooling involves facing the selection problem. The problem arises because the youth who choose to enroll in school are those who expect schooling to have favorable outcomes for them. If expected outcomes are related to objective ones, then the outcomes experienced by youth who choose to enroll in school differ from those that nonenrollees would experience if they were to enroll. Likewise, the outcomes experienced by nonenrollees differ from those that enrollees would experience if they were not to enroll. See, for example, Griliches (1977), Heckman and Robb (1985), and Manski (1989).

The selection problem implies that any effort to infer the objective returns to schooling from observations of realized outcomes requires at least some knowledge of the way youth make their schooling decisions. But we have already observed that, lacking data on the expectations of youth, one can only learn youth's decision rules conditional on maintained assumptions on expectations. Hence, one can only infer the objective returns to schooling conditional on the validity of expectations assumptions.

It is important to understand that these identification problems arise even in a stationary world, where the objective returns to schooling are constant over time. This will be illustrated through an example in section 2.3. Further identification problems may arise in a world with aggregate productivity shocks, where the objective returns to schooling change with time.

2.1.3   The Econometrics of Expectations Formation

The two identification problems just described would not be of concern if there were reason to think that prevailing expectations assumptions are correct. Logic and some indirect empirical evidence suggest otherwise. In particular, there is little reason to think that all youth form their expectations in the same way.

The logical point is that youth forming expectations face the same kind of inferential problem as do econometricians measuring educational productivity. Youth and econometricians may possess different data on realized outcomes, may have different knowledge of the economy, and may process their information in different ways. But both want to use their data and knowledge to learn the objective returns to schooling conditional on the available information. It follows that youth, like econometricians, face the selection problem. If youth use data on realized outcomes to form their expectations, then their interpretation of these data must depend on how they think other youth make schooling decisions. Expectations formation will be homogeneous only if all youth make the same assumptions about the behavior of their peers.

The empirical evidence is indirect but, I believe, compelling. Although we lack data on the expectations of youth, we have extensive data on the practices of econometricians studying educational productivity. For 30 years, in perhaps hundreds of published studies, econometricians have sought to learn the objective returns to schooling. Reading this literature reveals that econometric studies of the returns to schooling vary greatly in the conditioning variables used, in the outcome data analyzed, and in their handling of the selection problem. Compare, for example, Willis and Rosen (1979) and Murphy and Welch (1989). The former study analyzes data from the NBER-Thorndike Survey, estimates returns to schooling conditional on measured ability, and is explicitly concerned with the effect of unmeasured ability on the selection of students into schooling. The latter piece analyzes data from the Current Population Surveys, which contain no ability measures, and implicitly assumes that the selection of students into schooling is unrelated to ability. If experts can vary so widely in the way they infer the returns to schooling, it is reasonable to suspect that youth do, as well.

2.1.4   Elaboration on the Basic Ideas

The remaining sections of this paper elaborate on the foregoing basic ideas. Section 2.2 indicates that, if economists want to learn how youth perceive the returns to schooling, we cannot rely on the expectations research performed by other social scientists. Section 2.3 uses a simple formal model to show the different patterns of choices and outcomes that can result if youth do or do not condition their expectations on ability. Section 2.4 makes concluding comments on expectations research in economics.

2.2   Expectations Research in Psychology and Sociology

In contrast to economists, psychologists and sociologists routinely collect and analyze subjective data of many kinds, including expectations data from youth. I have sought to determine whether useful lessons can be extracted from these literatures. Unfortunately, my findings have been largely negative.

2.2.1   Measurement of Expectations

The prevailing measurement practice is to interpret responses to loosely worded questionnaire items as indicators of youths' expectations. Berndt and Miller (1990), for example, ask their sample of junior high school students to respond, on a five-point scale, to the question "How valuable do you think your education will be in getting the job you want?" Mickelson (1990) asks her sample of high school seniors to express their degree of agreement with the statement "Studying in school rarely pays off later with good jobs." Most of the literature poses such vague questions. An exception is a recent study of the income expectations of college seniors, by Smith and Powell (1990). These authors ask respondents to make unconditional forecasts of their "anticipated annual income in 10 years" and their "expected earnings" in the first

year of their first job. They also ask respondents to provide similar forecasts for the average member of their class.

2.2.2 Theories of Expectation Formation

The looseness with which psychologists and sociologists measure youth's expectations is matched by looseness in their thinking about expectations formation. Researchers in these fields theorize verbally rather than mathematically. As a consequence, it is even difficult to determine whether different researchers interpret the term *expectations* in a common, coherent fashion.[2]

The central social psychological idea is that expectations formation is a social phenomenon, each person learning about his prospects by observing the experiences of others. Bandura (1986, 47) writes:

> If knowledge could be acquired only through the effects of one's own actions, the process of cognitive and social development would be greatly retarded. . . . Fortunately, most human behavior is learned by observation through modeling. By observing others, one forms rules of behavior, and on future occasions this coded information serves as a guide for action. . . . Much social learning is fostered by observing the actual performances of others and the consequences for them.

This statement seems sensible; indeed I could interpret it as endorsing the idea that youth learning the returns to schooling are implicit econometricians. Unfortunately, the social psychological literature does not go much beyond the generalities expressed by Bandura. A long line of research, beginning with Hyman (1942), has sought to operationalize the idea that individuals learn from their "reference groups"; Bank, Slavings, and Biddle (1990) give an interesting historical account. But the idea of a reference group seems as amorphous today as it was 50 years ago.

It appears to me that if social psychologists are to make progress in understanding expectations formation, they must end their dependence on verbal reasoning, which invites conceptual ambiguity and logical inconsistency. Coherent analysis of complex social processes demands the discipline of formal modeling.

2.3 A Model of Information, Schooling Choices, and Outcomes

I observed in section 2.1 that some econometric studies (e.g., Willis and Rosen 1979; Manski and Wise 1983) assume that youth condition their expectations on their ability, while other studies (e.g., Freeman 1971; Murphy and Welch 1989) assume that they do not. Given the variation in econometric

---

2. There are mathematical psychologists who interpret expectations in the same subjective probabilistic way as economists do. See, for example, Kahneman and Tversky (1979) or Camerer and Kunreuther (1989). Their work, however, seems to have had no impact on psychologists or sociologists concerned with schooling behavior.

practice, it is of interest to determine how observed patterns of schooling choices and outcomes may depend on this aspect of expectations.

To address the question, I pose a simple stationary human capital model (section 2.3.1) and consider two alternative assumptions on expectations: myopic youth either condition expectations on ability (assumption A) or they do not (assumption B). I then derive the schooling choices and outcomes that result in the two cases (section 2.3.2). It turns out that in both cases there is a unique equilibrium in which expectations, although myopic, are fulfilled. But the characteristics of these equilibria differ. The main findings are:

- Assumption A yields a rational expectation equilibrium. Assumption B yields equilibrium expectations which are fulfilled, yet systematically incorrect.
- Fewer low-ability and more high-ability youth enroll under expectations assumption A than under B.
- The gross enrollment rate under A may be less or greater than under B, depending on the values of the model parameters.
- For some parameter values, the mean income realized by enrollees is known to be higher under assumption A than under B.

Having compared the two patterns of choices and outcomes, I consider the implications of misspecifying expectations for econometric analysis of schooling behavior (section 2.3.3). It is found that if youth do not condition their expectations on ability, then an econometrician who assumes they do so may mistakenly conclude from observed schooling behavior that youth are unconcerned with the returns to schooling.

2.3.1   The Model

Maintained Assumptions

Assume an overlapping-generations world in which each person lives for two periods. In the first period, a youth can choose to work ($c = w$) or to enroll in school ($c = s$); in the second period, all adults work. At the time of the schooling decision, youth know their real-valued ability z, their real-valued taste for schooling v, and the present discounted life-cycle log-income η that they would receive if they were to work immediately; for simplicity, assume that η is constant across the population and normalize the income scale by setting $\eta = 0$. Youth do not know the discounted log-income y they would receive if they were to enroll in school; y is a random variable whose realization becomes known after schooling is completed.

Each youth's value of (y, v, z) is independently drawn according to the following time-stationary process:

$$y = \alpha_1 + \beta_1 z + \epsilon_1, \quad \beta_1 \geq 0$$
(1)
$$v = \alpha_2 + \beta_2 z + \epsilon_2$$

$$(\epsilon_1, \epsilon_2, z) \sim N \left(0, \begin{array}{ccc} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & 1 \end{array} \right)$$

Thus, the objective probability distribution of (y, v, z) is trivariate normal. Letting z be a standard normal random variable and assuming that $\beta_1 \geq 0$ are normalizations that make ability a well-defined concept. Assuming the variance of $(\epsilon_1, \epsilon_2, z)$ to be diagonal is a real restriction; conditional on ability z, a youth's postschool income y and taste for schooling v are statistically independent.

The youth in a given generation share certain information about the schooling choices and realized incomes of the preceding generation. Let $E^*(y \mid z, v)$ be a youth's subjective expected value of y conditional on (z, v) and the common information. The decision rule is

$$(2) \qquad \begin{aligned} c &= s \text{ if } E^*(y \mid z, v) + v > 0 \\ &= w \text{ otherwise} . \end{aligned}$$

Expectations Assumptions

The model is complete when the subjective expected income $E^*(y \mid z, v)$ is specified. Although I have earlier criticized the prevailing assumption that expectations formation is homogeneous, I retain that assumption here.

The recent fashion in economics has been to assume that expectations are rational; youth a priori know that equation (1) holds and so set

$$(3) \qquad E^*(y \mid z, v) = \alpha_1 + \beta_1 z .$$

The realism of this assumption is most questionable. Having witnessed the struggles of econometricians to learn the returns to schooling, I find it difficult to accept the proposition that adolescents are endowed with this knowledge.

I instead assume that youth form their expectations in the manner of practicing econometricians: youth observe the incomes realized by members of the preceding generation who chose schooling, and they make inferences from these observations. But what information do they possess about the experiences of the preceding generation, and how do they use this information to form their expectations? I shall consider two cases of myopic expectations. In each case a youth, having observed the mean income $E_0(y \mid \Omega, c = s)$ realized by those members of the preceding generation who chose schooling and who had specified characteristics Ω, believes that he or she will receive the same mean income.[3] The two cases differ in the characteristics Ω on which youth condition their expectations. They are

---

3. The mean income $E_0(y \mid \Omega, c = s)$ is well defined only if there exist members of the preceding generation who chose schooling and who had characteristics Ω. The assumptions made in this section guarantee that this condition is satisfied (see Manski 1991).

$$\text{Assumption A: } E^{*A}(y \mid z, v) = E_0(y \mid z, c = s)$$
$$\text{Assumption B: } E^{*B}(y \mid z, v) = E_0(y \mid c = s).$$

Youth might form expectations as in assumption A if they observe the abilities and realized incomes of those members of the preceding generation who chose schooling. Suppose, however, that youth cannot observe the abilities of their elders. Unaware that income varies with ability, they might then form expectations as in assumption B.[4]

2.3.2   Schooling Choices and Realized Incomes

The two expectations assumptions imply systematically different patterns of schooling choices and realized incomes. To see this, I first derive the choice and income patterns that emerge under the two assumptions.

Expectations Conditioned on Ability and Schooling

By (2), a youth's schooling choice c is a function of his or her ability-taste pair (z, v). By (1), income y is statistically independent of v, conditional on z. Hence,

$$(4) \qquad E_0(y \mid z, c = s) = E_0(y \mid z) = \alpha_1 + \beta_1 z.$$

Thus, in the time-stationary environment (1), the myopic expectations (A) turn out to be rational. (These expectations would not generally be rational if the process generating (y, v, z) were not time-stationary.)

By (1), (2), and (4), the decision rule is

$$(5) \qquad \begin{aligned} c &= s \text{ if } \alpha_1 + \alpha_2 + (\beta_1 + \beta_2) z + \epsilon_2 > 0 \\ &= w \text{ otherwise.} \end{aligned}$$

So the probability that a youth with ability z selects school is

$$(6) \qquad P_A(c = s \mid z) = \Phi\left[\frac{\alpha_1 + \alpha_2 + (\beta_1 + \beta_2) z}{\sigma_2}\right],$$

where $\Phi(.)$ is the standard normal distribution function. The unconditional probability of schooling is

$$(7) \qquad P_A(c = s) = \Phi(\gamma_A),$$

where $\gamma_A \equiv (\alpha_1 + \alpha_2)[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2}$.

The mean income realized by youth with ability z who choose schooling is

---

4. Other specifications for Ω may be of interest. For example, Streufert (1991) assumes that youth observe the abilities, choices, and incomes of residents of their neighborhoods. He also supposes that neighborhoods are segregated by income classes. These assumptions suggest the expectations model

$$E^*(y \mid z, v) = E_0[y \mid z, y \epsilon(a, b), c = s] ,$$

where $[a, b]$ is the interval of incomes found in a youth's neighborhood.

$\alpha_1 + \beta_1 z$. Thus, income expectations are fulfilled. The mean income realized by all youth who choose schooling is

$$
\begin{aligned}
E_A(y \mid c=s) &= E[y \mid \alpha_1 + \alpha_2 + (\beta_1 + \beta_2)\, z + \epsilon_2 > 0] \\
&= \alpha_1 + \beta_1\, E[z \mid \alpha_1 + \alpha_2 + (\beta_1 + \beta_2)\, z + \epsilon_2 > 0] \\
&= \alpha_1 + \frac{\delta_A \phi(\gamma_A)}{\Phi(\gamma_A)},
\end{aligned}
$$
(8)

where $\delta_A \equiv \beta_1(\beta_1 + \beta_2)\,[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2}$ and $\phi(.)$ is the standard normal density function.

Expectations Conditioned on Schooling Only

Suppose that assumption B holds. Then the decision rule is

$$
\begin{aligned}
c &= s \quad \text{if } E_0\,(y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0 \\
&= w \quad \text{otherwise.}
\end{aligned}
$$
(9)

So the probability that a youth with ability z selects school is

$$
(10) \qquad P_B(c=s \mid z) = \Phi\left[\frac{E_0(y \mid c=s) + \alpha_2 + \beta_2 z}{\sigma_2}\right],
$$

and the unconditional probability of schooling is

$$
(11) \qquad\qquad\qquad P_B(c=s) = \Phi(\gamma_B),
$$

where $\gamma_B \equiv [E_0(y \mid c=s) + \alpha_2](\beta_2^2 + \sigma_2^2)^{-1/2}$.

The mean income realized by youth with ability z who choose schooling remains $\alpha_1 + \beta_1 z$ as before. The mean income realized by all youth who choose school is

$$
\begin{aligned}
E_B(y \mid E_0\,(y \mid c=s) &= E[y \mid E_0(y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0] \\
&= \alpha_1 + \beta_1 E[z \mid E_0\,(y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0] \\
&= \alpha_1 + \frac{\delta_B \phi(\gamma_B)}{\Phi(\gamma_B)}
\end{aligned}
$$
(12)

where $\delta_B \equiv \beta_1 \beta_2(\beta_2^2 + \sigma_2^2)^{-1/2}$.

Suppose, as seems reasonable, that the taste for schooling does not decrease with ability; that is, let $\beta_2 \geq 0$. Then there is a unique $E_0(y \mid c=s) \geq \alpha_1$ such that expectations are fulfilled. To see this, observe that expectations are fulfilled if (12) holds with $E_B(y \mid c=s) = E_0\,(y \mid c=s)$; that is, if

$$
(13) \quad E_0(y \mid c=s) = \alpha_1 + \beta_1 E[z \mid E_0(y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0].
$$

If $\beta_2 = 0$, (13) is solved at $E_0(y \mid c=s) = \alpha_1$. If $\beta_2 > 0$, (13) is solved at some $E_0(y \mid c=s) > \alpha_1$; this is so because $E[z \mid E_0\,(y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0]$ is a differentiable, strictly decreasing function of $E_0\,(y \mid c=s)$ whose value falls to 0 as $E_0\,(y \mid c=s)$ rises.

Observe that equilibrium expectations under assumption B, even though fulfilled, are systematically incorrect except in the special case $\beta_2 = 0$. Unconditional on ability, a youth's *objective* expected income following schooling is α_1. But it has just been shown that, in equilibrium, youth's common *subjective* expected income exceeds α_1 whenever $\beta_2 > 0$.

The fulfilled-expectations equilibrium (13) is globally stable when $\beta_1 < \beta_2$; I do not know the stability properties when $\beta_1 \geq \beta_2$. To show that $\beta_1 < \beta_2$ implies global stability, observe that global stability is guaranteed if the derivative of the right-hand side of (13) with respect to $E_0 (y \mid c=s)$ is always less than one in absolute value. It is shown in Goldberger (1983) that $0 < \partial E(z \mid z<t)/\partial t < 1$ for all real t; hence $-1 < \partial E(z \mid z>-t)/\partial t < 0$. It follows that, for all $[E_0 (y \mid c=s), \epsilon_2]$,

$$(14) \quad -\beta_1/\beta_2 < \beta_1 \frac{\partial E [z \mid E_0 (y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0, \epsilon_2]}{\partial E_0 (y \mid c=s)} < 0 .$$

Taking the expectation over ϵ_2 of the derivative in (14) yields

$$(15) \quad -\beta_1/\beta_2 < \beta_1 \frac{\partial E [z \mid E_0 (y \mid c=s) + \alpha_2 + \beta_2 z + \epsilon_2 > 0]}{\partial E_0 (y \mid c=s)} < 0 .$$

So the derivative is less than one in absolute value if $\beta_1 < \beta_2$.

Comparative Schooling Choices

The remainder of this section compares the patterns of schooling choices and realized incomes that emerge under the two expectations assumptions. In this discussion, I assume that the taste for schooling does not decrease with ability; that is, $\beta_2 \geq 0$. In discussing expectations assumption B, I restrict attention to the fulfilled-expectations equilibrium (13).[5]

Let us first compare the ability-conditioned enrollment probabilities $P_A(c=s \mid z)$ and $P_B(c=s \mid z)$, given in (6) and (10). Recall that the solution to (13) is $E_0(y \mid c=s) = \alpha_1$ if $\beta_2 = 0$, and satisfies $E_0(y \mid c=s) > \alpha_1$ if $\beta_2 > 0$. Hence, evaluated at $z = 0$,

$$(16) \quad \begin{aligned} P_A(c=s \mid z=0) &= P_B(c=s \mid z=0) \text{ if } \beta_2 = 0 \\ P_A(c=s \mid z=0) &< P_B(c=s \mid z=0) \text{ if } \beta_2 > 0 . \end{aligned}$$

This and the fact that $(\beta_1 + \beta_2) > \beta_2$ imply that

$$(17) \quad P_A(c=s \mid z) < P_B(c=s \mid z), \text{ all } z < 0 .$$

On the other hand, (16) and the fact that $(\beta_1 + \beta_2) > \beta_2$ imply that there exists a $z_0 \geq 0$ such that

$$(18) \quad P_A(c=s \mid z) > P_B(c=s \mid z), \text{ all } z > z_0 .$$

---

5. Thus, this discussion is not concerned with the dynamic adjustment questions studied by Freeman (1971).

Thus, fewer low-ability youth and more high-ability youth enroll under expectations assumption A than under B.

Overall, enrollments under assumption A may be less or greater than under B, depending on whether γ_A is less or greater than γ_B (see equations 7 and 11). We find that

(19)
$$
\begin{aligned}
\gamma_B < \gamma_A < 0 \quad &\text{if } \alpha_1 + \alpha_2 < 0 \text{ and } \beta_2 = 0 \\
\gamma_A < \min(0, \gamma_B) \quad &\text{if } \alpha_1 + \alpha_2 < 0 \text{ and } \beta_2 >> \beta_1 \\
\gamma_A = \gamma_B = 0 \quad &\text{if } \alpha_1 + \alpha_2 = 0 \text{ and } \beta_2 = 0 \\
\gamma_A = 0 < \gamma_B \quad &\text{if } \alpha_1 + \alpha_2 = 0 \text{ and } \beta_2 > 0 \\
0 < \gamma_A < \gamma_B \quad &\text{if } \alpha_1 + \alpha_2 > 0 .
\end{aligned}
$$

Hence,

(20)
$$
\begin{aligned}
P_B(c=s) < P_A(c=s) < 1/2 \quad &\text{if } \alpha_1 + \alpha_2 < 0 \text{ and } \beta_2 = 0 \\
P_A(c=s) < \min[1/2, P_B(c=s)] \quad &\text{if } \alpha_1 + \alpha_2 < 0 \text{ and } \beta_2 >> \beta_1 \\
P_A(c=s) = P_B(c=s) = 1/2 \quad &\text{if } \alpha_1 + \alpha_2 = 0 \text{ and } \beta_2 = 0 \\
P_A(c=s) = 1/2 < P_B(c=s) \quad &\text{if } \alpha_1 + \alpha_2 = 0 \text{ and } \beta_2 > 0 \\
1/2 < P_A(c=s) < P_B(c=s) \quad &\text{if } \alpha_1 + \alpha_2 > 0 .
\end{aligned}
$$

If $\alpha_1 + \alpha_2 < 0$ and if β_2 and β_1 are the same order of magnitude, then the ordering of $P_B(c=s)$ and $P_A(c=s)$ appears to depend on the specific values of the model parameters.

Comparative Realized Incomes

The mean income realized by a youth of ability z who enrolls in school is $\alpha_1 + \beta_1 z$, whether expectations assumption A or B holds. The mean income of all enrollees depends on the ability distribution of enrollees and so varies with the expectations assumption, as follows. By (8) and (12),

(21)
$$
E_A(y \mid c=s) - E_B(y \mid c=s) = \frac{\delta_A \phi(\gamma_A)}{\Phi(\gamma_A)} - \frac{\delta_B \phi(\gamma_B)}{\Phi(\gamma_B)} .
$$

It can be shown that $\delta_A > \delta_B$ for all values of the model parameters; moreover, $\delta_B = 0$ if $\beta_2 = 0$.[6] The Mills ratio $\phi(.)/\Phi(.)$ is strictly decreasing in its argument, so

6. To prove that $\delta_A > \delta_B$, observe that

$$
\delta_A - \delta_B = \beta_1 \{ (\beta_1 + \beta_2)[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2} - \beta_2(\beta_2^2 + \sigma_2^2)^{-1/2} \} .
$$

$\beta_1 = 0 \implies \delta_A - \delta_B = 0$. The expression $(\beta_1 + \beta_2)[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2}$ increases with β_1, as

$$
\begin{aligned}
&\frac{\partial(\beta_1 + \beta_2)[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2}}{\partial \beta_1} \\
&= [(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2} - (\beta_1 + \beta_2)^2[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-3/2} \\
&= [(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1/2} \{ 1 - (\beta_1 + \beta_2)^2[(\beta_1 + \beta_2)^2 + \sigma_2^2]^{-1} \} \\
&> 0 .
\end{aligned}
$$

Hence $\beta_1 > 0 \implies \delta_A - \delta_B > 0$.

(22) $\gamma_A \leq \gamma_B \Rightarrow E_A(y \mid c=s) > E_B(y \mid c=s).$

Hence, by (19),

(23) $E_A(y \mid c=s) > E_B(y \mid c=s) \begin{cases} \text{if } \alpha_1 + \alpha_2 < 0, \beta_2 >> \beta_1 \\ \text{if } \alpha_1 + \alpha_2 \geq 0 \,. \end{cases}$

Equation (23) shows that, for some values of the model parameters, the mean realized income of school enrollees is higher under expectations assumption A than under B. I have not been able to determine the relationship between $E_A(y \mid c=s)$ and $E_B(y \mid c=s)$ for other parameter values.

2.3.3   Econometric Analysis with Misspecified Expectations

Analysis of Behavior

It remains to inquire into the consequences for econometric analysis of misspecifying expectations. Consider the following idealized description of an econometric analysis of schooling choices: For each member of a random sample of youth, an econometrician observes (c, z) and observes y when $c = s$; he does not observe v. The econometrician assumes that (1) describes the objective probability distribution of (y, v, z) and that (2) is the decision rule youth use to make their schooling choices. As is common in the literature, he assumes that tastes for schooling are independent of ability; that is, $\beta_2 = 0$. Moreover, he makes the conventional assumption that expectations are rational.

Believing that (6) describes choice behavior and that $\beta_2 = 0$, the econometrician would form the probit model

(24) $P[c=s \mid E^*(y \mid z, v) = \alpha_1 + \beta_1 z] = \Phi(\pi_0 + \pi_1 z)$

and estimate (π_0, π_1) by maximum likelihood. He would interpret π_0 to be $(\alpha_1 + \alpha_2)/\sigma_2$ and π_1 to be β_1/σ_2.

Suppose that the econometrician is correct in assuming (1) and (2) but incorrect otherwise; in fact, β_2 may be positive and assumption B holds. Then (10) describes actual choice behavior, and the econometrician's interpretation of (π_0, π_1) is incorrect. In reality, $\pi_0 = [E_0(y \mid c=s) + \alpha_2]/\sigma_2$ and $\pi_1 = \beta_2/\sigma_2$.

The misinterpretation of π_1 is of particular interest. The econometrician believes π_1 to measure educational decisions' sensitivity to changes in the income returns to schooling. In fact, π_1 measures the degree to which tastes for schooling vary with ability. Suppose, for example, that $\beta_2 = 0$ as assumed. Then $\pi_1 = 0$. Finding this, the econometrician would conclude that, in making their schooling choices, youth are unconcerned with the income returns to schooling. This conclusion would, of course, be incorrect. If the returns to schooling were to shift through a change in α_1, then the intercept π_0 would change and so would the probability of enrolling.

Analysis of the Returns to Schooling

I have made the idealized assumption that the econometrician observes ability z without error. Given this, data on enrolled youths' abilities and realized incomes can be used to obtain a consistent least-squares estimate for the parameters (α_1, β_1). It might therefore seem that, if z is observed, analysis of the objective returns to schooling requires no knowledge of how youth make their schooling choices. But there is an implicit expectational assumption, namely that youth do not know ϵ_1 at the time of their schooling decisions. If this assumption fails, then the econometrician's estimate of (α_1, β_1) is not consistent. The selection problem implies that an econometrician analyzing the returns to schooling must take a stand on the information youth use in forming their expectations.

2.4   Conclusion: Expectations Research in Economics

The question posed in the title of this paper cannot be answered at this time. Having chosen to make assumptions rather than to investigate expectations formation, economists do not know how youth infer the returns to schooling. If youth form their expectations in anything like the manner that econometricians study the returns to schooling, then prevailing expectations assumptions cannot be correct. Without an understanding of expectations, it is not possible to interpret schooling behavior nor to measure the objective returns to schooling. As a consequence, the economics of education is at an impasse.

As I see it, progress is possible only if economists become more willing to entertain the use of subjective data in empirical analysis. Decisions under uncertainty reflect the interplay of preferences, expectations, and opportunities. Choice data alone cannot disentangle these factors. The identification problem can be solved if choice data are combined with interpretable subjective data on expectations and/or preferences.

The question, of course, is whether interpretable subjective data can be obtained. The dominant view expressed by economists today is negative. In particular, economists often assert that respondents to surveys have no incentive to answer questions carefully or honestly; hence, they conclude, there is no reason to think that subjective responses reliably reflect respondents' thinking. But this reasoning is not applied consistently. Empirical economic analyses of schooling behavior routinely use respondents' self-reports of their backgrounds, choices, and outcomes. Many analyses use scores on tests administered with surveys to measure respondents' ability. Thus, ironically, economists' own revealed preferences in empirical analysis are somewhat at variance with their expressed views about the interpretability of survey data.

It should be noted that economists' views on the use of subjective data have not always been so negative. In the 1940s it was common to interview businessmen about their expectations and decision rules. In an influential article,

Machlup (1946) sharply attacked then-existing survey practices as not yielding credible information. This article apparently played an important role in dampening the enthusiasm of economists for subjective data. But Machlup only sought to criticize the collection of subjective data through standardized questionnaires. He stressed that cost and revenue expectations are subjective. He advocated research in which the economist learns the institutional peculiarities of a firm and then questions its managers in language they understand.

From the mid-1950s through the mid-1960s, economists analyzed data on consumers' buying intentions (see, e.g., Juster 1966). Although this practice has since almost ceased among economists, it remains firmly entrenched among demographers and market researchers. I have recently reviewed and reinterpreted this literature in Manski (1990).

The early literatures on businessmen's expectations and on consumers' intentions may hold lessons for efforts to learn youth's expectations. The present problem, however, seems more difficult than those treated previously. Whereas past efforts have sought to elicit unconditional forecasts from adult respondents, here we need to elicit choice-conditioned forecasts from adolescent respondents. We shall not know whether this is feasible until we try.

References

Bandura, A. 1986. *Social foundations of thought and action.* Englewood Cliffs, N.J.: Prentice-Hall.

Bank, B., R. Slavings, and B. Biddle. 1990. Effects of peer, faculty, and parental influences on students' persistence. *Sociology of Education* 63:208–25.

Berndt, T., and K. Miller. 1990. Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology* 82:319–26.

Camerer, C., and H. Kunreuther. 1989. Decision processes for low probability events: Policy implications. *Journal of Policy Analysis and Management* 8:565–92.

Freeman, R. 1971. *The market for college-trained manpower.* Cambridge, Mass.: Harvard University Press.

Goldberger, A. 1983. Abnormal selection bias. In *Studies in econometrics, time series, and multivariate statistics,* ed. S. Karlin, T. Amemiya, and L. Goodman. Orlando, Fla.: Academic Press.

Griliches, Z. 1977. Estimating the return to schooling: A progress report. *Econometrica* 45:1–22.

Heckman, J., and R. Robb. 1985. Alternative methods for evaluating the impact of interventions. In *Longitudinal analysis of labor market data,* ed. J. Heckman and B. Singer. Cambridge: Cambridge University Press.

Hyman, H. 1942. The psychology of status. *Archives of Psychology,* no. 269.

Juster, T. 1966. Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association* 61:658–96.

Kahneman, D., and A. Tversky. 1979. Prospect theory. *Econometrica* 47:263–92.

Machlup, F. 1946. Marginal analysis and empirical research. *American Economic Review* 36:519–54.

Manski, C. 1989. Anatomy of the selection problem. *Journal of Human Resources* 24:343–60.

———. 1990. The use of intentions data to predict behavior: A best case analysis. *Journal of the American Statistical Association* 85:934–40.

———. 1991. Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty. In *Nonparametric and semiparametric methods in econometrics and statistics,* ed. W. Barnett, J. Powell, and G. Tauchen. New York: Cambridge University Press.

Manski, C., and D. Wise. 1983. *College choice in America.* Cambridge, Mass.: Harvard University Press.

Mickelson, R. 1990. The attitude-achievement paradox among black adolescents. *Sociology of Education* 63:44–61.

Murphy, K., and F. Welch. 1989. Wage premiums for college graduates: Recent growth and possible explanations. *Educational Researcher* 18(4):17–26.

Smith, H., and B. Powell. 1990. Great expectations: Variations in income expectations among college seniors. *Sociology of Education* 63:194–207.

Streufert, P. 1991. The effect of underclass social isolation on schooling choice. Institute for Research on Poverty Discussion Paper no. 954-91. University of Wisconsin–Madison.

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

Comment   Eric A. Hanushek

Charles F. Manski has a history of contributions to the understanding of higher education and, particularly, college choice that goes back farther than that of virtually everybody participating in this conference. Moreover, his contributions have been especially important, bringing serious analytical effort to bear on an area that tends to be punctuated more by fuzziness. Thus, it is good to have him return to the general area.

His paper pursues an extraordinarily important set of issues: How do prospective students form expectations about the advantages of higher education? How do expectations condition reality and the outcomes that are observed? And how do the efforts of analysts interact with the actual choice process of students?

The overall idea is quite straightforward. If students' expectations determine the pattern of college enrollment, ignoring expectations could lead to selection problems that imply biased estimates of the value of schooling. Unfortunately, however, economists have directed virtually no attention to understanding the expectations problem in higher education. Manski is led to conclude that the only hope for understanding not only college choice but also labor market returns on schooling is the collection and analysis of subjective data on students' expectations.

Eric A. Hanushek is professor of economics and political science at the University of Rochester.

I am sympathetic with his concerns about the need to understand expectations of students. I think that this is an important area of research. Investment in human capital is really an exercise in decision making under uncertainty, and the character of individual perceptions and expectations must be important. Moreover, having just participated in the college decision-making process as a parent, I am struck by how far assumptions about complete information appear to be from reality. Finally, I am very supportive of more refinements in modeling college choice in general. My caricature of the current state of modeling is that the analysis begins with a standard human capital investment model which compares expected benefits to the costs of schooling but then, in the empirical work, turns to a simple regression of college attendance on tuition. While there are clear exceptions, much of the work in this general area is simply very primitive. Therefore, the systematic study of this by Manski is most welcome. Of course, as critic, I must also be clear that I am less persuaded that nothing can be done about understanding the returns to schooling or the choice of colleges without delving into subjective views of students. And I am currently unsure of exactly how Manski would have us proceed.

The paper has two distinct sections. The first argues the general principle that understanding expectations is important, while the second works through a specific model. I will consider each in turn.

The overall motivation for considering student expectations follows as a logical extension of much of the current work in labor economics. The discussion of income determination has been dominated for some 30 years by consideration of unmeasured individual characteristics and how neglect of these might bias statistical results. Manski simply takes this argument a step further: if trained econometricians have such difficulties, surely high school seniors also have problems. A central feature is then to understand how different expectations of students affect actual outcomes.

The general discussion of expectations makes two points. First, expectation formation is central to much of current economics, both micro and macro, but the underlying basis for individual expectations is often not even discussed, let alone analyzed. Manski highlights the importance of individual expectations about future earnings, an appropriate starting point; but the issues are clearly much larger, including such things as expected schooling costs and attrition probabilities. Second, other disciplines which purportedly consider expectations—psychology and sociology—do not do very well at it.

Frankly, this motivation does not lead me to great optimism about our ability to generate or use subjective expectations information in refining our ideas of income determination or college choice. For some time, economists have flirted with the idea of using subjective information for predictive or interpretative purposes. I think, for example, of debates in the 1950s about the importance of expectations in determining investment decisions in productive capacity. While there continue to be periodic surveys of the sentiments of

purchasing agents and the like, I see little evidence that such subjective expectations information has made great inroads into aggregate econometric models or that it has helped in producing improved forecasts of investment activity. This is matched by specialists in other disciplines who purport to be able to measure expectations but do not appear able to do so.

The implicit argument in Manski's paper, moving to the second general section, is that developing an explicit model of behavior will inform us on what data to collect and what subjective information to gather. It will also inform us on how subjective information affects specification and estimation of the choice model. Unfortunately, I think his formal models tend to confuse the issues and to distort the analysis.

The basic model has three distinct features: formation of expectations about future earnings, unmeasured individual heterogeneity (which, in the time-honored labor economics tradition, is simply labeled "ability"), and heterogeneity of individual taste for schooling. The unfortunate part of his specific model is that the results depend crucially on the full structure of the model. While similar results might come from other models, much of the leverage of this structure relates directly to individual tastes for schooling—something that is quite independent of student expectations or how they are formed. To be clear, allowing for heterogeneity of tastes is not inherently peculiar, but it does make the expectations story very hard to parse out. Moreover, many of the central results in the theoretical section appear to evaporate if tastes do not enter the model systematically. In other words, as I work through the formal models, the role of expectations does not seem to be central to the results, even though the paper starts and ends with a plea for better understanding of expectations.

At the outset, I had hoped that the Manski model and foray into expectation formation might shed light on some currently perplexing issues about college attendance. Specifically, I was hoping that the explicit consideration of expectations might aid in untangling the effects of the substantial changes in relative wages on college-going behavior. From the mid-1970s to the mid-1980s, the wage premium for college education (as opposed to high school education) appeared to explode. By the estimates of Murphy and Welch (1989)—which Manski discounts because of lack of consideration of expectations—the college premium went from roughly 30 percent to 70 percent for new labor market entrants. How have students taken this information into account, and does expectations formation explain the apparently sluggish response of students to what appear to be extremely strong market forces? Or, on a slightly different front, is there something about income expectations that fits into the apparently peculiar pattern of relative black and white college attendance in the 1980s? These issues, while alluded to in Manski's discussion, are entirely different from the ones he considers—the role of ability and tastes in a simple selection model where all individuals of the same ability face the same income stream. When I look at the aggregate raw data, I suspect that the effects of

overall college–high school earnings differences and changes in the differential are much more important than pure ability rents and misperceptions of ability rents.

Let me return to the starting point. I could not agree more with Manski that student expectations must be central to the college choice problem. I also believe that we possess a very primitive understanding of expectations, even though we have elevated the role of expectations in virtually every area of economics. Finally, I believe that economists have much to offer in measuring and understanding expectations, because the empirical force of expectations can only be understood from an underlying decision theoretic perspective. On the other hand, my comments should suggest that it might be some time before investments in understanding expectations from subjective data pay off. I conclude that we should pursue better measurement and analysis of expectations. I also conclude that, short of this, there are useful things that economists can do to understand better college choice, income determination, and the like.

Reference

Murphy, Kevin, and Finis Welch. 1989. Wage premiums for college graduates: Recent growth and possible explanations. *Educational Researcher* 18(4):17–26.