University of Wisconsin-Madison

Institute for Research on Poverty

Discussion Papers

Charles F. Manski

SCHOOLING AS EXPERIMENTATION:
A Reappraisal of the College
Dropout Phenomenon

DP #865-88

Institute for Research on Poverty
Discussion Paper no. 865-88

SCHOOLING AS EXPERIMENTATION:
A REAPPRAISAL OF THE COLLEGE DROPOUT PHENOMENON

Charles F. Manski
Department of Economics
and
Institute for Research on Poverty
University of Wisconsin-Madison

September 1988

This research was partially supported under a contract from the National Assessment of Vocational Education to MPR Associates, Inc. I have benefited from the comments of Arthur Goldberger and Norton Grubb.

ABSTRACT

College dropout is widely considered a social problem. In fact, reducing dropout would not necessarily make society better off. This conclusion derives from analysis of the process of college enrollment and completion. The key observation is that students contemplating college entrance do not know whether completion will be feasible or desirable. Hence, enrollment is a decision to initiate an experiment, one of whose possible outcomes is dropout. Experiments should be evaluated by their ex ante expected return, not by their ex post success rate. It follows that, told only the completion rate of enrolled college students, one cannot judge whether the right enrollment decisions have been made.

1. INTRODUCTION

College dropout is widely considered a social problem. Reducing dropout is often cited as an objective of student financial aid. For example, in a recent issue of Change, Fischer (1987) says: "All knowledgeable observers bemoan current dropout levels and believe society would be better off if these levels were lower" (p. 42). The presumption that current dropout levels are too high leads Fischer to propose that the existing system of grants and loans to students enrolled in postsecondary education be replaced by a "graduation-contingent" aid program.

Writing in the Economics of Education Review, James (1988) states: "And, does this aid accomplish one of its major purposes, reducing the above-average attrition rates of low socio-economic status (SES) students, so that more of them complete college?" (p. 3). In the same issue, Stampen and Cabrera (1988) examine the effect of student aid on dropout and provide numerous references to previous work on the subject.

I shall argue here that the conventional wisdom regarding college dropout has no normative basis. Lowering dropout levels would not necessarily make society better off. Student aid policy should not be evaluated by its effect on dropout.

2. SCHOOLING AS EXPERIMENTATION

Suppose that a student has enrolled in college. What determines whether he or she will obtain a degree? College completion presumably requires that two conditions hold. First, the student must be able to pass the prescribed courses. Second, the student must decide that it is worthwhile to persist to graduation. Thus, college completion has both exogenous and endogenous determinants. The student must be able to graduate and must want to.

Now consider a student contemplating enrollment. At this point, the student does not know whether he has the ability to complete college. Nor does he know whether he will find it worthwhile to do so. The only way the student can definitively determine whether college is for him is by enrolling. Thus, the decision to enroll is a decision to initiate an experiment.

Viewing schooling as experimentation has important implications for our interpretation of college dropout. The usefulness of an experiment cannot be judged by its outcome. Nor does it suffice to know the experiment's ex ante probability of success. The appropriate way to evaluate an experiment is by its ex ante expected return.

An observer given data on the success rate of those experiments that are performed cannot, in general, judge whether the right experimentation decisions have been made. To see why, suppose that each member of a population must decide whether or not to perform an experiment which has cost C and, if successful, benefit B. Suppose that the probability of success, denoted P, varies across the population while C and B do not. Then an individual should perform the experiment if his expected return PB - C exceeds zero. That is, the threshold success probability at which experimentation becomes worthwhile is C/B. The success rate of those experiments that are performed is $E(P|P>C/B)$, the expected value of P within the subpopulation for whom P exceeds C/B. This rate depends on the manner in which P varies across the population and on the value C/B. In principle, it can be any number between zero and one.

This reasoning applies to the analysis of enrollment and persistence. Let schooling have cost C and, if completed successfully, benefit B. Let P be the probability of completion. Suppose that an observer is told the completion rate of enrolled students but is not told C, B, nor the manner in which P varies across students. Then the observer cannot judge whether the right enrollment decisions have been made.

College dropout and high school dropout are fundamentally different phenomena. College enrollment is voluntary; high school enrollment is compulsory. A student entering college recognizes that dropout may be the outcome and feels it worthwhile to accept this risk. A student entering high school does not thereby signal his acceptance of the risk of dropout. Many high school dropouts are people who, in the absence of compulsory attendance laws, would have chosen not to enroll.

The observation that postsecondary schooling is an experiment is not new. Manski and Wise (1983) note:

> "Like trial and error in the job market, postsecondary education may for many young people be part of the search process that leads to discovery of what they like and don't like and of which occupations are compatible with their interests and abilities. To this extent, students may derive informational value from attendance, even if they drop out." (p. 10)

Fischer (1987) says: "There are so many college dropouts for the same reason there are so many small business failures - start-up costs are not exorbitant and the risk is rationally worth taking" (p. 44).

Curiously, the implications of thinking of schooling as experimentation seem not to have been worked out. In particular, it has not been appreciated that college dropout statistics carry no normative message. Much of the literature on the economics of education ignores dropout entirely by treating schooling as an investment which, when undertaken, will definitely be completed. Recent survey articles by Blaug (1985) and Freeman (1986) make no mention of the dropout phenomenon.

Those studies which treat schooling as an investment with uncertain outcome do not analyze the interaction between ex ante dropout probabilities, enrollment decisions, and subsequent dropout levels. See, for example, Chapters 6 and 8 of Manski and Wise (1983). One study, by Comay, Melnik, and Pollatschek

(1973), does present a model which can be applied to study schooling as experimentation. These authors do not, however, develop the experimentation theme.

In an attempt to shed light on the interaction between dropout probabilities, enrollments, and realized dropouts, I develop here a model of college enrollment and completion. The simpler version of this model assumes that, conditional on enrollment, completion is exogenous. A more general version makes completion partly endogenous. Working through the implications of this model makes clear that student aid policy should not be evaluated by its effect on dropout. Two findings are especially striking.

First, setting the policy goal to be reduction of dropout yields the perverse conclusion that it would be best to eliminate student aid entirely. The reason is that eliminating aid makes college enrollment less attractive relative to working. Lowering the attractiveness of enrollment reduces the number of students who choose to enroll. The students who choose to work rather than enroll are those with the lowest college completion probabilities. Hence, eliminating aid shifts the composition of enrollment towards those students with the highest completion probabilities.

Second, suppose it is possible to introduce a policy which reduces the ex ante dropout probability of each member of the population. This policy may either reduce or raise ex post dropout, for the following reason. A policy which reduces dropout probabilities does lower dropout among those students who would have enrolled in the pre-policy regime. But introduction of the policy also induces new students to enroll. The observed dropout level (i.e., the number of dropouts) rises if the number of induced enrollees who drop out exceeds the gain in college completion among existing enrollees. The dropout rate (i.e., the fraction of enrollees who drop out) rises if the completion probabilities of induced enrollees are sufficiently lower than those of existing enrollees.

The analysis supporting these findings is presented below.

3. MODELS OF COLLEGE ENROLLMENT AND COMPLETION

The word model sketched in Section 2 presumes that college completion is exogenous; the experiment either succeeds or fails. Section 3.1 develops a formal model elaborating on this idea. Sections 3.2 and 3.3 use the model to study the determination of aggregate college enrollment and completion. Section 3.4 extends the analysis to allow for the possibility that college completion is partly endogenous.

3.1. A Model with Exogenous Completion

Assume that a student graduating from high school may either work or enter college. Let V_w denote the expected utility of working, V_c the expected utility associated with completing college, and V_d the expected utility associated with dropping out. Let P denote the probability of completing college should the student enroll. Then the student will enroll if

$$(1)\quad PV_c + (1-P)V_d > V_w.$$

Note that the student is indifferent between enrolling and working if $PV_c + (1-P)V_d = V_w$. Provided that the number of students exactly on the margin is negligible, it is innocuous to assume that all such students enroll.

In principle, all of the quantities V_w, V_c, V_d, and P may vary across students. To make the main points, it is simplest to condition on specified values of the expected utilities (V_w, V_c, V_d) and to consider the population of students characterized by these values. These students may vary in their college completion probabilities P.

We shall focus on the case in which $V_d < V_w < V_c$. Otherwise the analysis is trivial. In particular, if $V_w < V_d < V_c$, then every student enrolls, regardless of his completion probability. If $V_d < V_c < V_w$, then no student enrolls. If $V_c < V_d$, then the enrolled students prefer to drop out rather than graduate. Given that $V_c > V_d$, the enrollment criterion (1) is equivalent to

$$(2)\quad P > \frac{V_w - V_d}{V_c - V_d}.$$

Let

$$(3) \quad \pi \equiv \frac{V_w - V_d}{V_c - V_d}$$

be the threshold completion probability at which enrollment becomes worthwhile. Let F denote the distribution of P across students, conditional on the specified values of V_w, V_c, and V_d. Let Q denote the college enrollment level, that is, the fraction of the student population who choose to enroll. Then Q is the fraction of the population for whom P exceeds π. That is,

$$(4) \quad Q = \int_{\pi}^{1} dF.$$

Let Q_c denote the college completion level, that is, the fraction of the population who enroll in and complete college. Then

$$(5) \quad Q_c = \int_{\pi}^{1} PdF.$$

Let Q_d denote the college dropout level, that is, the fraction of the population who enroll in college and drop out. Then

$$(6) \quad Q_d = Q - Q_c = \int_{\pi}^{1} (1-P)dF.$$

Finally, let R_d denote the college dropout rate, that is, the fraction of enrollees who drop out. Then

$$(7)\quad R_d = \frac{Q_d}{Q} = \frac{\int_{\pi}^{1}(1-P)\,dF}{\int_{\pi}^{1}dF}\ .$$

3.2. Effect of a Change in π

In this subsection and the next, we ask how the quantities Q, Q_c, Q_d and R_d are affected by changes in π and F.

Here we consider a rise in π, holding F fixed. By (3), a rise in π can be achieved by increasing the expected utility of working, by decreasing the expected utility of completing college, or by decreasing the expected utility of dropping out. That is, a rise in π follows from any change that makes college less attractive relative to working. Proposition 1 gives the qualitative consequences.

PROPOSITION 1: Suppose that the threshold completion probability at which enrollment becomes worthwhile rises from π to some $\lambda > \pi$. Then

A. The college enrollment level Q falls.

B. The college completion level Q_c falls.

C. The college dropout level Q_d falls.

D. The college dropout rate R_d falls. ∎

This proposition is proved in the Appendix. The reasoning can be explained easily. Raising π obviously reduces the number of students who choose to enroll (Part A). Hence it reduces the number who complete college (Part B) and the number who drop out (Part C). The students who choose to work rather than enroll are those with the lowest college completion probabilities. Hence, raising π shifts the composition of enrollment towards those students with the highest completion probabilities (Part D).

Proposition 1 shows why student aid policy should not be evaluated by its effect on college dropout. Suppose that a policy change worsens the terms of aid. Then, ceteris paribus, π rises. So the dropout level Q_d and the dropout rate R_d both fall. Thus, evaluating aid policy by its effect on college dropout yields the perverse conclusion that aid should be reduced to zero. What this conclusion ignores, of course, is that reducing aid lowers the college completion level as well.

As stated, Proposition 1 conditions on specified values of (V_c,V_d,V_w). That is, the Proposition concerns a population of students who have the same expected utility values but who vary in their college completion probabilities. Parts A through C hold unconditionally. If Q, Q_c, and Q_d fall conditional on every possible value of (V_c,V_d,V_w), then these quantities necessarily fall in the aggregate. Part D, which involves a rate rather than a level, need not hold unconditionally.

3.3. Effect of a Change in F

Consider now the effects of a change in F, holding π fixed. Many types of changes might be contemplated. We shall examine an especially simple case. Suppose that the college completion probability of each member of the population rises. This may, for example, be achieved by improving the quality of high school education or by providing tutoring while in college. Proposition 2 gives the qualitative consequences.

PROPOSITION 2: Suppose that each college completion probability P rises to some $g(P) > P$. Then

A. The college enrollment level Q rises.

B. The college completion level Q_c rises.

C. The college dropout level Q_d may rise or fall.

D. The college dropout rate R_d may rise or fall. ∎

Proposition 2 is proved in the Appendix. This proposition provides further evidence that policy should not be evaluated by its effect on college dropout. It might have been thought that a policy which raises all college completion probabilities must lower the level and rate of realized dropout. In retrospect, it is easy to see why this is not so.

A policy change which raises completion probabilities does lower dropout among students who enroll in the pre-change regime. But the change also induces new students to enroll (Part A). Of these new students, some complete college (Part B). Others do not. The aggregate dropout level rises/falls if the number of induced enrollees who drop out is larger/smaller than the reduction in dropout among the pre-change enrollees (Part C). The dropout rate rises if the college completion probabilities of induced enrollees are sufficiently lower than those of pre-change enrollees. Otherwise the dropout rate falls (Part D).

As stated, Proposition 2 conditions on specified values of (V_c, V_d, V_w). The entire proposition holds unconditionally. Parts A through C concern levels, so the reasoning applied to Proposition 1 applies here as well. Part D states that, conditional on (V_c, V_d, V_w), the rate R_d can either rise or fall. If so, then R_d can obviously either rise or fall unconditionally.

3.4. A Model with Partly Endogenous Completion

This section generalizes the foregoing analysis by making college completion partly endogenous.

Assume that an enrolled student completes college if he or she works hard enough. It may be that the effort needed to graduate is infinite, so that graduation is impossible. This is equivalent to saying that the student does not have the requisite ability. On the other hand, it may be that finite effort suffices. If so, the student decides whether exerting that effort is worthwhile. The student can determine the required effort only by enrolling. Before enrolling, he has effort expectations. In what follows, we first formalize the completion decision and then work backwards to the enrollment decision.

As earlier, let V_c be the expected utility associated with completing college and V_d be the expected utility associated with dropping out. Let $R \equiv V_c - V_d$. Let Z denote the effort required to graduate, a non-negative value expressed in units of utility. Then an enrolled student will choose to complete college if

(8) $Z < R$

and to drop out otherwise.

Now consider a student facing the enrollment decision. The student knows V_c, V_d, and V_w. Not yet having enrolled, he does not know Z. He believes, however, that Z will be drawn from some probability distribution G. In this setting, the expected utility of enrollment is

$$\int (V_c - Z)1[Z<R]dG + \int V_d 1[Z>R]dG$$

$$= V_c \mathrm{Prob}(Z<R) + V_d \mathrm{Prob}(Z>R) - \int_0^R Z dG.$$

Hence the student chooses to enroll if

(9) $$V_c \mathrm{Prob}(Z<R) + V_d \mathrm{Prob}(Z>R) - \int_0^R Z dG > V_w.$$

(This assumes that a student who is indifferent between enrolling and working does enroll.)

Conditioning on specified values for (V_c, V_d, V_w), the enrollment decision is determined by the student's effort expectations, as embodied in G. A particularly simple case is that in which G is Bernoulli, with probability P that Z = 0 and probability 1-P that $Z = \infty$. Here Prob(Z<R) = P, Prob(Z>R) = 1-P, and

$$\int_0^R Z dG = 0,$$

whatever non-negative value R may take. It follows that the completion probability for an enrolled student is P and that the enrollment criterion (9) reduces to

$$PV_c + (1-P)V_d > V_w.$$

Thus, making G Bernoulli with mass points at zero and infinity generates the model with exogenous completion of Section 3.1.

Propositions 1 and 2 hold for other specifications of G that make college completion partly endogenous. A complete analysis will not be attempted here. Instead, we shall consider a simple generalization of the Bernoulli model. Assume that, for each student, G is Bernoulli, with probability P on the event Z = K and 1-P on the event $Z = \infty$. The exogenous completion model made K = 0. Here K is a non-negative value that varies across students.

For this specification, Prob(Z<R) = P, Prob(Z>R) = 1-P, and

$$\int_0^R Z dG = K$$

if $K < R$. On the other hand, $\text{Prob}(Z<R) = 0$, $\text{Prob}(Z>R) = 1$, and

$$\int_0^R Z dG = 0$$

if $K > R$. Hence the college completion probability for an enrolled student is P if $K < R$ and zero if $K > R$. The enrollment criterion is

$$(10) \quad 1[K<R]\{PV_c + (1-P)V_d - PK\} + 1[K>R]V_d > V_w.$$

Provided that $V_d < V_w$, (10) is equivalent to saying that the student enrolls if both of the following conditions hold:

$$(11a) \quad K < R$$

$$(11b) \quad P > \frac{V_w - V_d}{R - K} .$$

Thus, students for whom $K > R$ do not enroll in college. A student for whom $K < R$ enrolls if his probability of completion exceeds the threshold

$$(12) \quad \pi_K = \frac{V_w - V_d}{R - K} .$$

The enrollment criterion (11b) has the same form as the criterion (2) that applies when completion is exogenous. We may therefore conclude that, conditioning on specified values for (V_c, V_d, V_w) and K, Propositions 1 and 2 hold.

4. CONCLUSION

The simple analysis of this paper suffices to show that dropout statistics per se carry no normative message. This conclusion ultimately derives from two simple observations. First, college enrollment is voluntary, not compulsory. Second, the decision to enroll is a decision to initiate an experiment, a possible outcome of which is dropout. Hence, college enrollments and completions are jointly determined.

One should not interpret the foregoing as saying that present rates of college enrollment and completion are necessarily socially optimal. Consideration of one scenario suffices to demonstrate how social and private interests may diverge.

Let the assumptions of Section 3.4 hold. Suppose that society values college completion more than students do privately. Suppose that the social and private values of working and dropping out coincide. Then the socially optimal enrollment criterion is given not by (11a)-(11b) but rather by

(13a) $K < R + G$

(13b) $P > \dfrac{V_w - V_d}{R + G - K}$,

where G is the positive difference between the social and private values of college completion. Comparing (11) and (13), we see that society prefers a higher enrollment level than that generated privately. In particular, society prefers that a student enroll if his completion probability is above the threshold in (13b). This threshold is lower than the private one given in (11b). Hence, in this scenario, society prefers a higher college dropout rate than that generated privately.

APPENDIX: PROOF OF PROPOSITIONS 1 AND 2

PROOF OF PROPOSITION 1: Parts A, B, and C follow immediately from equations (4), (5), and (6) respectively. To prove Part D, first define

$$\alpha \equiv \int_{\pi}^{\lambda} dF \qquad \beta \equiv \int_{\lambda}^{1} P\,dF \qquad \gamma \equiv \int_{\lambda}^{1} dF$$

and note that $\lambda\gamma < \beta$. Next observe that, by (7), the dropout rate under λ minus that under π is

$$\frac{\int_{\lambda}^{1}(1-P)\,dF}{\int_{\lambda}^{1} dF} - \frac{\int_{\pi}^{1}(1-P)\,dF}{\int_{\pi}^{1} dF} = -\frac{\int_{\lambda}^{1} P\,dF}{\int_{\lambda}^{1} dF} + \frac{\int_{\pi}^{\lambda} P\,dF + \int_{\lambda}^{1} P\,dF}{\int_{\pi}^{\lambda} dF + \int_{\lambda}^{1} dF}$$

$$< -\frac{\int_{\lambda}^{1} P\,dF}{\int_{\lambda}^{1} dF} + \frac{\lambda\int_{\pi}^{\lambda} dF + \int_{\lambda}^{1} P\,dF}{\int_{\pi}^{\lambda} dF + \int_{\lambda}^{1} dF} \equiv -\frac{\beta}{\gamma} + \frac{\lambda\alpha + \beta}{\alpha + \gamma}$$

$$= \frac{-\beta\alpha - \beta\gamma + \lambda\alpha\gamma + \beta\gamma}{\gamma(\alpha+\gamma)} = \frac{\alpha(\lambda\gamma-\beta)}{\gamma(\alpha+\gamma)}\,.$$

We noted earlier that $\lambda\gamma < \beta$. Hence the above expression cannot be positive.

Q.E.D.

PROOF OF PROPOSITION 2: In what follows, we use 1[ ] to denote the indicator function taking the value one if the logical event inside the brackets is true and zero otherwise. In particular, $1[g(P)>\pi] = 1$ if $g(P) > \pi$ and $1[g(P)>\pi] = 0$ if $g(P) < \pi$.

Part A: By (4), the enrollment level before the change is $\int_{\pi}^{1} dF$. Following the change it is

$$\int_{\pi}^{1} dF + \int_{0}^{\pi} 1[g(P)>\pi]dF \quad > \quad \int_{\pi}^{1} dF.$$

Part B: By (5), the completion level before the change is $\int_{\pi}^{1} PdF$. Following the change, it is

$$\int_{\pi}^{1} g(P)dF + \int_{0}^{\pi} g(P)1[g(P)>\pi]dF \quad > \quad \int_{\pi}^{1} PdF.$$

Part C: By (6), the dropout level before the change is $\int_{\pi}^{1} (1-P)dF$. Following the change, it is

$$\int_{\pi}^{1} [1-g(P)]dF + \int_{0}^{\pi} [1-g(P)]1[g(P)>\pi]dF.$$

Depending on $g(*)$, the post-change dropout level can be either higher or lower than the pre-change one. To show this, it suffices to consider two special cases.

Consider first any $g(*)$ such that $g(P) < \pi$ for $P < \pi$. Then the post-change dropout level is

$$\int_{\pi}^{1}[1-g(P)]dF \quad < \quad \int_{\pi}^{1}(1-P)dF.$$

Next consider any g(*) such that $g(P) = P$ for $P > \pi$. Here the post-change dropout level is

$$\int_{\pi}^{1}(1-P)dF + \int_{0}^{\pi}[1-g(P)]1[g(P)>\pi]dF \quad > \quad \int_{\pi}^{1}(1-P)dF.$$

Part D: By (7), the dropout rate before the change is

$$\frac{\int_{\pi}^{1}(1-P)dF}{\int_{\pi}^{1}dF} \; .$$

Following the change, it is

$$\frac{\int_{\pi}^{1}[1-g(P)]dF + \int_{0}^{\pi}[1-g(P)]1[g(P)>\pi]dF}{\int_{\pi}^{1}dF + \int_{0}^{\pi}1[g(P)>\pi]dF} \; .$$

Depending on g(*), the post-change dropout rate can be either higher or lower than the pre-change one. To show this, it again suffices to consider two special cases.

Consider first any g(*) such that $g(P) < \pi$ for $P < \pi$. Then the post-change dropout rate is

$$\frac{\int_{\pi}^{1}[1-g(P)]dF}{\int_{\pi}^{1}dF} < \frac{\int_{\pi}^{1}(1-P)dF}{\int_{\pi}^{1}dF} .$$

Next consider the particular transformation $g(P) = \pi$ for $P < \pi$ and $g(P) = P$ for $P > \pi$. Here the post-change dropout rate is

$$\frac{\int_{\pi}^{1}(1-P)dF + (1-\pi)\int_{0}^{\pi}dF}{\int_{\pi}^{1}dF + \int_{0}^{\pi}dF} = \int_{\pi}^{1}(1-P)dF + (1-\pi)\int_{0}^{\pi}dF$$

$$= \frac{\int_{\pi}^{1}(1-P)dF}{\int_{\pi}^{1}dF} \int_{\pi}^{1}dF + (1-\pi)\int_{0}^{\pi}dF.$$

That is, the post-change dropout rate is a weighted average of the pre-change rate and of $(1-\pi)$. The pre-change rate is smaller than $(1-\pi)$. Hence, the post-change rate is larger than the pre-change one.

Q.E.D.

References

Blaug, M. (1985), "Where Are We Now in the Economics of Education?" Economics of Education Review, 4, 17-28.

Comay, Y., Melnik, A. and Pollatschek, M. (1973), "The Option Value of Education and the Optimal Path for Investment in Human Capital," International Economic Review, 14, 421-435.

Freeman, R. (1986), "Demand For Education," in O. Ashenfelter and R. Layard (Eds.), Handbook of Labor Economics, Amsterdam: Elsevier.

Fischer, F. (1987), "Graduation-Contingent Student Aid," Change, November/December, 40-47.

James, E. (1988), "Student Aid and College Attendance: Where Are We Now and Where Do We Go from Here?" Economics of Education Review, 7, 1-14.

Manski, C. and Wise, D. (1983), College Choice in America, Cambridge: Harvard University Press.

Stampen, J. and Cabrera, A. (1988), "The Targeting and Packaging of Student Aid and Its Effect on Attrition," Economics of Education Review, 7, 29-46.