

DETERMINANTS AND DISTRIBUTIONAL ASPECTS OF ENROLLMENT IN U.S. HIGHER EDUCATION*

ARTHUR J. CORAZZINI
DENNIS J. DUGAN
HENRY G. GRABOWSKI

ABSTRACT

Based upon human capital theory, an enrollment model for higher education is formulated with demand being subject to nonprice rationing by academic admission standards. Cross-sectional differences in student enrollment are related to variables representing both demand factors and supply-side constraints. Two questionnaire surveys — Project Talent's national cross-sectional sample in the early 1960s and a recent survey of 4,000 high school seniors in the Boston SMSA — provide sufficient data to test the theoretical hypothesis derived. At both levels of aggregation, strong structural relationships between college attendance and socioeconomic status emerge. Stratifying the on-going group by socioeconomic quartiles yields insights into the distributional aspects of higher education enrollment.

I. INTRODUCTION

Many would consider U.S. higher education the growth industry of the 1950s and 1960s. By several measures, the higher education industry grew rapidly

The authors are, respectively, Assistant Professor of Economics, Tufts University; Associate Professor of Economics, University of Notre Dame; and Assistant Professor of Economics, Yale University.

- * This national cross-sectional analysis, with evidence from one urban area, is an outgrowth of a joint research effort carried out with Ernest Bartell, C.S.C., John Keith, and Alvin Klevorick for the State of Massachusetts Board of Higher Education. The authors wish to thank all three gentlemen for numerous suggestions and helpful comments. They also wish to thank the State of Massachusetts for granting the funds which made this research possible.

in both absolute and relative terms during these two decades.¹ This growth is not surprising when one considers the high priority that the public has traditionally placed upon the role of higher education in our society. In addition to being thought of as a powerful vehicle of socioeconomic mobility and a major determinant of social stability, it is a common contention that higher education produces a pool of highly skilled manpower needed to support the technological progress of our society and that the general welfare is increased especially because of the knowledge produced in the higher education sector.

The Demand for Higher Educational Services

Consider the hypothetical situation in which a high school graduate is faced with an "either-or" decision of going on to college or entering the labor force. If the student follows rational investment decision criteria, he will decide to go on to college if the present value of the benefits associated with going on (discounted at the appropriate rate) are at least equal to the present discounted value of both the direct and the opportunity costs of doing so. The present value of the benefits can be divided into two basic components: the expected value of the stream of increased earnings that accrue from a college education, and the value of any direct consumption benefits from undertaking this activity. The direct cost of going to college will be the sum of several charges such as tuition, extraordinary living costs at college, and special educational fees. The opportunity costs, in turn, will be directly related to the wage that could be earned at the best job alternatives during the time spent in attendance at college.

These individual cost/benefit calculations provide us with a 0–1 variable of an individual's enrollment decision as a function of a number of determining factors. The aggregate number of high school students in a region who decide to attend college will also depend upon representative group values of the determinant factors affecting the individual's decision. Thus, a formal statement of demand that includes both investment and consumption aspects of higher education is given by the following equation:

1 For example, in 1951 the number of students enrolled in institutions of higher education as a percent of the population 18–24 years of age was 13.4, but by 1968, 30.3 of those in that age bracket were enrolled in such institutions. In absolute terms, enrollment growth in the industry is impressive. In 1950, 2.66 million people were enrolled in higher education, and by 1969, the figure stood at 7.1 million. Furthermore, the growth in higher education has not been limited to increases in the size of student bodies. In dollar terms, the growth in the industry is even more definite. Expenditures per student for the 1949–50 academic year were \$850, compared to \$2,257 for 1965–66 (in current dollars), the last year for which expenditure data are available.

$$(1) \quad Q_i = f(r_i, y_i, z_i, p_i, c_i, n_i)$$

where Q_i is the number of high school graduates in region i who wish to attend college, y_i is the expected economic value in terms of increased earnings in region i resulting from higher education, z_i is the expected economic value from the direct consumption benefits resulting from higher education in region i , p_i are the direct costs of the higher educational services in region i , c_i are the indirect or opportunity costs of higher educational services in region i , r_i is a representative discount rate for the group, and n_i is the eligible population in region i .

If it is assumed that the demand function is homogeneous of degree one in n_i , equation (1) can be converted into the general form:

$$(2) \quad Q_i/n_i = D_i = f(r_i, y_i, z_i, p_i, c_i)$$

Equation (2) implies that changes in n_i across regions do not carry with them compositional changes in the population of eligibles that might affect Q_i . Thus, the dependent variable D_i is the percentage of students in state i finding it economically desirable to enroll in college. This percentage approximates the probability that a high school student selected randomly in a given region will find it economically advantageous to go to college.

A general rise in the expected benefits, either in the stream of earnings or consumption benefits, should increase the percentage of students finding it economically desirable to pursue higher education. An increase in the cost of education investments, either in the form of increased direct educational charges or in the form of an increase in the opportunity cost of attending, should lead to a reduction in the enrollment decision. Finally, because most of the benefits of higher education are future in nature as opposed to the costs which are predominantly current, an increase in the discount rates should also lead to a reduction in the demand for college enrollment. If the present is deemed much more important than the future by a high discount rate, the direct cost of education, and especially the opportunity cost of attending, will loom as the most important factors to the decision maker.

Supply-Side Constraints

The higher education market is supplied by private nonprofit institutions as well as publicly supported colleges and universities. The prices charged for higher educational services are subsidized to varying degrees by public and private sources. Moreover, a considerable amount of price discrimination exists in the form of scholarships and differential tuition fees. The spectrum of existing prices is generated by a complex administrative process which is influenced by economic and noneconomic factors.

For publicly supported institutions, pricing decisions are in the hands of public officials such as state executives and legislators. Very often, an

admission acceptance policy is established at a fixed nominal price for all students who are above some target admission standard, generally based on high school grades, academic ranking, and the Scholastic Aptitude Tests (SAT). The setting of admission standards is often a more crucial element of public policy than the price charged to students, which is usually nominal in nature.

For private colleges and universities, prices also have an administered character. However, the prices charged to students are more strongly tied to cost considerations, and there is undoubtedly a strong element of cost push in the determination of their administered price. That is, by habit or through necessity, a certain target level or proportion of the revenues at private schools are traditionally raised through student payments. Nationally known schools, where demand is oversubscribed, appear to take the lead in this price-setting process, and the other private institutions, some of which have excess supply, come into some kind of equilibrium relationship with them.

If the above description is correct, it suggests that prices in higher education will be largely supply-side determined. In addition, it suggests that at the prevailing level of prices some of the aggregate demand will be constrained by admission standards. Hence, corresponding to the demand function formulated above, a rationing function may be specified to reflect the effect of admission policies in rationing demand. In distinguishing the effect of these policies on various high school populations, particularly relevant is the performance of the group vis-à-vis the standards employed in the particular "market" where their demand is registered. Accordingly, the determinants of rationing are represented in function form as:

$$(3) \quad R_i = g(A_i, S_{1i}, S_{2i})$$

where R_i is the percent of the eligible population in region i who desire college but are unable to obtain admission, A_i is an index of region i 's population aptitude for college as evaluated by college officials, S_{1i} is an index of admission standards employed at publicly supported institutions in the region, and S_{2i} is an index of the admission standards employed at the private institutions relevant to the particular high school population. The nature of this rationing of higher education spaces is such that an increase in the population's aptitude should lead to a reduction in R_i , whereas an increase in admissions standards should lead to an increase in the number who are unable to obtain admission.

Determinants of Enrollment

Now the enrollment percentages observed from various populations of eligible students will be positively related to factors increasing the potential

demand from each group and negatively related to those factors constraining or rationing this demand. In particular, given the way that we have defined the above variables, we have the mathematical relation or identity that:

$$(4) \quad E = D - R$$

In the next section we proceed to develop some empirical counterparts for the variables influencing D and R as given in equations (2) and (3) as a basis for estimating an enrollment function E presented in equation (4). As one might expect, it has not been possible to construct empirical variables which distinguish all of these factors separately. In general, the cost variables can be distinguished quite well from other effects. However, a few proxy variables will be used to encompass jointly several of the noncost variables operating here. Consequently, one cannot identify directly the degree to which enrollment is separately affected by demand-side factors and supply-side constraints. Nevertheless, the regression analysis using these joint proxy variables gives some insight on the analytical question of interest here—the distributional impact of these variables on various socioeconomic groups. Moreover, some direct evidence on the effect of particular supply-side constraints in one urban area will be presented in a later section.

II. THE NATIONAL CROSS-SECTION REGRESSION ANALYSIS

The empirical measures available as proxies for the explanatory variables in equations (2) and (3) include tuition rates at various types of higher education institutions, state labor market variables of wage rates and unemployment rates, average level of father's education, and high school student's performance on achievement tests.² Average tuition fees at public

2 The empirical analysis presented in this section relies on data from Project Talent's national sample of high school students in the early 1960s. Each individual in this sample, as a 10th grader, was given a common battery of achievement tests and interviewed regarding his socioeconomic background. These data were subsequently collated with information on whether the student enrolled in college in the year following graduation of his high school class. The data from this national sample were cross-sectionally stratified by the student's state of residence during high school, and average values were computed for each state subsample. These subsample means, together with direct and opportunity cost variables constructed at the state level, serve as basic inputs for the regression analysis. Two types of variables that were presented in theoretical discussions are not included in the empirical analysis because satisfactory empirical counterparts for these variables were not available. Specifically, there are no variables reflecting the level of state scholarship and loan program aid and the academic admission standards at the publicly supported institutions in each state. However, the availability of national NDEA loans in 1963 would have the tendency of equalizing the opportunity of students in the various states to obtain loan aid, thus minimizing the effect of interstate differences

four-year universities, private universities, junior colleges, and teacher colleges are introduced to represent direct charges of higher education in a state. Although tuition charges do not encompass all the direct costs of college, they do reflect a large portion of total outlay of students on tuition, books, fees, and differential living expenses.³

The empirical proxies for the opportunity cost of attending college are the state average wage rate for production and nonsupervisory workers and the level of unemployment. The latter indicates the probability of finding employment if one is in the labor force.⁴ The level of father's education represents several theoretical variables. Because it is positively related to family income, paternal education is an appropriate index of the student's ability to finance college education. It also serves to index the intellectual environment experienced by a student in his home life and, therefore, will reflect personal tastes and preferences for a college education and perhaps expected return from college as well. Student performance and achievement tests reflect not only a student's ability to overcome the rationing that may exist in college admission policies but also will be related to his preference

in loans upon enrollment rates. The effect of omitting these variables may introduce some bias into the coefficient estimates, particularly those of the educational price variables. Of course, to the extent that omitted variables are not strongly correlated with the other independent variables, they will be absorbed into the error term and merely increase the size of the unexplained variance. On a priori grounds, however, one cannot reject the possibility that state university and college systems tend to follow parallel policies with regard to price, admission standards, and the availability of scholarship funds. If this were so, omission of the latter two factors in the regression analysis would lead to an overestimate of the price effect, for the price variable would be capturing the combined effect of other types of educational rationing.

- 3 Because students are free to attend college out of state and a significant number actually do, it may seem inappropriate to use in-state tuition prices as the relevant costs of higher education for high school populations in each state. Nevertheless, significant in-state admission and price discrimination exist at the public institutions in each state, producing a distinct asymmetry in educational costs favoring in-state attendance in public universities. Moreover, transportation costs necessarily also produce *de facto* price discrimination for students attending colleges and universities at distances from their homes. While in-state educational prices are certainly not the relevant costs for all students, they are the relevant costs for the majority of students. Therefore, they could provide a good first approximation to the direct costs of education facing the different student populations.
- 4 Statewide averages are employed for these variables, and the possibility of out-of-state migration in seeking job opportunities makes this only an approximation of the true opportunity cost. Furthermore, the state unemployment rate may not be as appropriate as the teenage (14 to 19 years of age) unemployment rate for each state. The teenage rate is not available for 1963, but the 1960 rate is and it is highly correlated to the total unemployment rate with a simple correlation of .79.

and tastes for education and probably his expected return from a college education.

The higher education enrollment function may now be expressed in terms of the empirical variables. Its general functional notation is:

$$(5) \quad E_i = H(T_{ji}, T_{ui}, T_{ci}, T_{pi}, W_i, U_i, F_i, A_i)$$

where E_i is the percent of 10th grade (1960) high school students in state i who enrolled in college in 1963; T_{ji} , T_{ui} , T_{ci} , T_{pi} are state average tuition rates at junior colleges, public four-year universities, teacher colleges, and private four-year universities, respectively; W_i is average earnings of production workers in state i ; F_i is average level of father's education in state i ; and A_i is ability as measured by performance on achievement tests. The enrollment function is expressed in linear form to be empirically estimated by linear regression analysis:

$$(6) \quad E_i = A_0 + a_1T_{ji} + a_2T_{ui} + a_3T_{ci} + a_4T_{pi} + a_5W_i + a_6U_i \\ + a_7F_i + a_8A_i + e_i$$

where e_i is a statistical error term.

There is good reason to believe that these enrollment relationships are nonlinear. That is, an increase in average family income of a given amount in the lowest income quartile would likely increase the enrollment rate for that group by a greater amount than the change in enrollment for the highest income quartile, given a similar change in its family income variable. To compensate for this possible distortion, the sample was stratified by income (socioeconomic status) and separate regressions were run for each quartile. This disaggregated analysis yields insights into the distributional effects of the explanatory variables on various socioeconomic groups.

III. ESTIMATION OF ENROLLMENT FUNCTION: NATIONAL RESULTS

The statistical results are presented in Table 1 for the analysis of a total enrollment function and for the socioeconomic quartile enrollment functions. In brief, the total enrollment function tests our hypotheses: (1) that the price and opportunity cost effect is negative, (2) that father's education has a positive effect, and (3) that achievement has a positive impact on enrollment. The empirical results of the quartile enrollment functions clearly exhibit the differential effects of price, father's education, and scholastic achievement upon different groups of students. Not only are the determinants of stratified groups of students significantly different in some cases, but also the impact of common explanatory factors vary widely for the enrollment rates of different quartiles.

TABLE 1
REGRESSION RESULTS OF HIGHER EDUCATION ENROLLMENT RATES
(IN PERCENT TERMS) OF 1960 TENTH GRADERS: TOTAL AND BY SOCIOECONOMIC QUANTILES

| Percent Enrolled | Constant | T_j | T_u | T_c | T_p | W | U | F | A | \bar{R}^2 |
|------------------|----------|-------------------|------------------|-----------------|------------------|------------------|------------------|------------------|------------------|-------------|
| E_T | 14.43 | -.011 (3.14)** | -.027 (2.32)* | -.005 (1.26) | -.009 (2.06)* | -3.62 (1.03) | .834 (2.04)* | 2.84 (4.21)** | .176 (5.72)** | .769 |
| E_{1SES} | 6.29 | -.028 (2.06)* | -.050 (1.71)* | -.024 (1.04) | .007 (0.57) | -5.39 (2.29)* | 3.104 (1.34) | 5.07 (2.52)** | .164 (1.56) | .537 |
| E_{2SES} | 16.62 | -.037 (1.49) | -.055 (1.81)* | .029 (0.96) | -.006 (1.09) | -6.36 (1.08) | 1.264 (1.22) | 3.41 (2.17)* | .212 (2.02)* | .412 |
| E_{3SES} | 23.65 | -.021 (1.19) | -.039 (1.91)* | -.007 (1.36) | -.018 (1.74)* | -4.01 (0.96) | -1.260 (0.58) | 3.63 (2.08)* | .276 (2.12)* | .450 |
| E_{HSES} | 22.36 | .002 (0.57) | -.050 (1.68)* | .002 (0.86) | -.016 (2.09)* | 2.75 (0.54) | -.260 (1.89)* | 1.07 (1.21) | .180 (2.23)* | .491 |

Notation: t -statistics are in parentheses. * = statistically significant at the 5 percent level (one-tailed test); ** = statistically significant at the 1 percent level (one-tailed test).

\bar{R}^2 = coefficient of determination adjusted for degree of freedom; E_T = total enrollment rate in college of 1960 10th graders (in percent); E_{1SES} = enrollment rate of low socioeconomic quartile; E_{2SES} = enrollment rate of second quartile; E_{3SES} = enrollment rate of third quartile; E_{HSES} = enrollment rate of high socioeconomic quartile; T_j = tuition at junior colleges; T_u = tuition at four-year public universities; T_c = tuition at teacher colleges; T_p = tuition at four-year private universities; W = average hourly earnings of production workers; U = unemployment rate; F = paternal education ranked by educational attainment groups of 0-7 years, 8 years, 1-3 years of high school, high school degree, 1-3 years of college, college degree, 5+ years of college; A = performance on Project Talent achievement tests.

Considering the total enrollment function first, it is clear that tuition and unemployment, empirical counterparts of the price variable, are statistically significant determinants of total enrollment. The total enrollment rate is most responsive to tuition charges at four-year public universities, and a decrease of \$100 in tuition in 1963 is associated with a 2.65 percent increase in the nation's enrollment, based upon these cross-section results. Junior college and private university tuition rates are also significant, but the magnitude of their impact upon enrollment is less than one-half that of tuition at four-year universities. The unemployment rate is a significant determinant of total enrollment, and it has the expected positive impact upon total enrollment. Hence, high unemployment rates within a state act as a definite deterrent to immediate entrance into the labor force. The other labor market variable, average hourly earnings, has a coefficient with the expected negative sign, suggesting that as the potential earnings of high school students increase, they are more inclined to participate in the labor force, and the enrollment rate in college falls.

The education of a student's father is also an important determinant of college attendance. A change of one unit in the educational attainment of the father is associated with an increase in total enrollments of 2.84 percent.⁵

As expected in the total enrollment function, the performance of students on achievement tests had a positive effect upon enrollment rates. As stated in the earlier theoretical discussion (Section II), the positive relationship measures the impact of two separate forces: (1) a taste for education on the part of the student as indicated by his ability to score well on the achievement tests which is an indication of how he currently performs as a student, and (2) an indication of the willingness of institutions of higher education to accept the student according to its standards. A high score on the Scholastic Aptitude Test ensures a place at a college or university, for that is indeed one of their rationing devices. Both of these forces, one a demand factor and the other a supply phenomenon, jointly act under one empirical, catch-all variable of achievement scores.

5 Associated with higher level of educational attainment are higher incomes, and a separate calculation indicates that a change of one unit (near the mean, in this case "1 to 3 years of college") in the educational attainment of the father is associated with a change in the father's income of \$2,381. Income of the head of the household, or even more appropriately family income, clearly represents a source of college funds to the high school student and therefore a possible bypass to the imperfections that exist in the human capital market. Furthermore, family income may embody the whole socioeconomic environment of the student which was the setting if not necessarily the cause of his level of scholastic achievement in high school. Therefore, the empirical relationship cited above does not establish a direct link between financial considerations and enrollment, for the paternal education variable has many facets, only one of which is financial ability.

Turning to the question of whether groups of different socioeconomic status are likely to react quite differently to the economic barriers to attendance at college, each state sample was stratified according to national socioeconomic quartile levels, and an enrollment equation for each quartile was empirically estimated. The adjusted coefficients of determination range from a high of .537 for the lowest quartile to .412 for the second lowest quartile; they are statistically significant. However, corresponding to these low coefficients of determination is a smaller number of statistically significant variables in each equation.⁶ The behavior of these variables, as a group, exhibits a good deal of plausibility. Tuition at public universities, which enroll the bulk of on-going students, is significant for all quartiles, whereas junior college tuition has its impact on the lower quartiles and private college tuition on the upper quartiles.

The empirical counterparts for the opportunity costs of higher education also have a differential impact on the enrollment rates of different quartiles. The regression coefficient of average hourly earnings is negative and highly significant for the lowest socioeconomic group. Correspondingly, the unemployment rate variable, while not statistically significant, has its largest impact on the lowest socioeconomic group. Moreover, the impact on college enrollment of these two variables on the second lowest socioeconomic quartile, while not statistically significant, is still quite large. Interestingly, the estimated coefficients in the upper income quartile for these two variables are opposite to what was expected on opportunity cost grounds. This may not be implausible and could reflect a substitution effect across socioeconomic classes via admission standards. That is, high opportunity costs cause more low-income students to forego college and take jobs in the labor market. This, in turn, makes it easier for students in higher socioeconomic groups, who are less sensitive to these kinds of opportunity costs, to obtain college enrollment. Whether this kind of mechanism is occurring or not, the overall pattern of results for the opportunity cost variables does indicate that high school students from low socioeconomic situations respond very definitely to the earnings they forego by not participating in the labor force, and that they are more likely to be diverted from pursuing higher education on this account than are their high socioeconomic counterparts.⁷

The paternal education and achievement test variables exhibit somewhat asymmetrical behavior across the four socioeconomic quartiles. The

6 This reduction in \bar{R}^2 is undoubtedly due in large part to the fact that by stratifying the sample in this fashion, the portion of the variation across quartiles is no longer present and the range of variation in each quartile is much smaller.

7 When the sample was stratified by male and female, the opportunity costs variables were significant for males but not for females.

paternal education variable has the expected positive sign for all four quartiles, but it has the largest effect on the lowest socioeconomic and the smallest on the upper socioeconomic quartile. On the other hand, the achievement score variable, while more uniform in magnitude, exhibits the opposite tendency. At first glance, these results would seem to indicate that it is variables for which paternal education is a proxy (that is, family background, motivation and preference for college enrollment, etc.) which have more of an impact on the enrollment of lower income students rather than their performance per se on admission tests and the variables for which it proxies. Aside from the possible problem of multicollinearity,⁸ a check on the data reveals that these results conceal a real issue. Specifically, relative to other quartiles, there is a lack of variation in the achievement test variable for the lowest quartile and for the paternal education variable in the highest quartile. Hence, a significant portion of the lowest quartile may be encountering the barrier of minimum admission requirements, and the current analysis would not have strong discriminatory power in showing it.⁹ Although the national results shed no light upon this possibility, empirical results from an analysis of a metropolitan area are quite relevant in this context.

Despite ambiguities in interpretation of this kind, the estimated coefficients for the paternal education and achievement score variables, taken together, indicate that the nonmarket factors for which they are proxies are significant determinants of college enrollment across the socioeconomic quartiles. For example, using the estimates in Table 1, an increase in the mean values of the paternal education and achievement test scores for the lowest quartile to the values realized by the second lowest quartile would result in an enrollment increase of over 5 percent. Similarly, a further increase up to the levels of the highest quartile would imply, using the estimated coefficient of the second and third quartiles as measures of the marginal impact, an additional increase on the order of 7 percent. Thus, these variables, along with price and cost factors, play a significant role in explaining differences in college enrollment rates across socioeconomic classes.

In summary, the empirical analysis of national enrollment rates demonstrates that there exists a substantial degree of responsiveness on the part of

8 While multicollinearity does not appear to be a severe problem here, the simple correlation coefficient between these two variables is on the order of 0.5 for the regressions in Table 1, and this is statistically significant at the 1 percent level.

9 Some indirect evidence for this position is provided from the results in Table 1 by the fact that the lowest socioeconomic class is particularly sensitive to the junior college tuition variable, and junior colleges traditionally have had the lowest admission standards.

the potential higher education to market and nonmarket factors. Furthermore, the results indicate that the effect of these factors are not uniform across all segments of the student population. After considering these issues further in the context of an analysis of one urban area—the Boston Standard Metropolitan Statistical Area (SMSA)—the policy implications of these results will be discussed in the final section.

IV. THE EMPIRICAL RESULTS: A LOCAL ANALYSIS

In order to gain some further insights into the distributional aspects of college enrollment and, in particular, rationing by college admission standards, we turn to an analysis of results from one urban locale—the Boston metropolitan area.¹⁰

The first issue that we shall consider is the variability in the number of students seeking college admission across high schools in the Boston SMSA. The state university and most private colleges require scores on the SAT as a prerequisite for admission. Therefore, the percentage of students who take the SAT may be used as an approximate index of the number of students entering the “market” for a college education from each SMSA high school.¹¹

Of course, from the standpoint of our previous discussion, this index will reflect not only differences in potential demand and motivation, but also

10 Most of the data here were obtained from a 10 percent structured sample of Boston SMSA high school students during the spring of 1969. These students provided information on college enrollment plans as well as on various aspects of their family background (father's occupation, maternal and paternal education, etc.). In addition, the scores of all Boston SMSA students on the Scholastic Aptitude Test (SAT) through December 1969 were made available to this study. While the data from this sample do not as yet allow us to say anything about actual enrollments, they do permit one to study various aspects of the admission process.

11 Some possible biases in using the measure in this fashion should be noted. First of all, some students who take the exam may never actually apply for college. That is, school or parental coercion may cause some students who do not wish to go to college to take the exam. However, data from the sample questionnaire of Boston SMSA indicate that relatively few students fall into this category. Only 120 students, or slightly less than 7 percent of the 1,782 students taking the SAT in the questionnaire sample, indicated no desire to go to college. Conversely, some students who seek to go to college may not take the SAT. Some colleges, particularly junior colleges, do not require the SAT for admission. However, our questionnaire sample indicated that virtually all students desiring to go to college had taken the SAT. This is consistent with the finding that only a small portion of the sample, approximately 7 percent, aspired to only junior college admission. Moreover, even the members of this group might find it necessary or advantageous to take the SAT exam.

ex ante rationing on the part of some students. That is, some students who desire college may have such pessimistic expectations concerning their ability to finance or be admitted to a college that they will not bother to even apply. Therefore, it at best captures only that component of demand for which student aspirations are sufficient to warrant the small investment in time and money normally associated with the application of admission for college.

Not surprisingly, the percentage of students seeking college admission by this measure varies greatly across the 56 public and private schools in the Boston SMSA which comprise our sample. Specifically, the number taking the SAT ranges in value from 9 percent in one inner city school to a full 100 percent for a few of the highly academically oriented schools in the Boston suburbs. As the high schools under study typically have a large number of students in the senior class (that is, more than 100), one can view these percentages as reasonable estimates of the *probability* that a student from a particular school will seek admission to an institution of higher education. The large observed variations in these percentages indicate considerable sample heterogeneity across schools with regard to the characteristics relevant to college aspirations.

Given our focus on the distributional aspects of college enrollment, we may inquire to what extent these observed percentage differences are an income-related phenomenon. That is, what is the *structural* relationship which exists in the Boston SMSA between a student seeking college and his family income? While we do not have any data on family income for individual students, we can obtain a reasonable estimate of the *mean* family income for the students enrolled in each high school by using data obtained in our 10 percent sample survey on the paternal occupation, education, and age (see Appendix for details). Hence, we can relate the observed percentages of students taking the SAT to estimates of mean family income derived in this way in order to get at the desired structural relationship.

Proceeding in this fashion, the following nonlinear regression equation was estimated, using a weighted regression procedure:¹²

$$(7) \quad S_j = -.59 + .180\bar{I}_j - .0023 \left[\sum_{i=1}^{n_j} (I_{ij}^2/n_j) \right] \\ (3.65) ** (0.74) \quad \bar{R}^2 = 0.54 \\ j = 1.56$$

12 Because the number of students varies significantly across schools, a weighted regression procedure was used in the empirical estimation of this equation. This involved weighting each observation (i.e., the means for each school) by the square root of the number of students in that school.

where S_j = the fraction of high school seniors in school j taking the SAT exam; \bar{I}_j = the sample estimate of the mean family income of high school seniors enrolled in school j (in thousands of dollars); n_j = the number of high school seniors in school j ; and $\sum_{i=1}^{n_j} (I_{ij}^2/n_j)$ = the sample estimate of the mean sum of squares of family income for seniors in high school j . The t -values are in parentheses.

The nonlinear formulation of equation (7), involving both terms in mean family income and the mean sum of squares of family income for each high school,¹³ was chosen to allow for the possibility of diminishing returns in this structural relationship. The negative coefficient on the sum of squares term does indicate the presence of diminishing returns, but this term has a relatively slight effect and is not statistically significant.¹⁴ On the other hand, the linear term is positive and highly significant. The strength of the structural tie between "demand" for higher education and family income is illustrated by the fact that this variable has an income elasticity of approximately 2 calculated at the sample means.

In analyzing the factors underlying this structural relation, we may refer once again to the theoretical framework of Section II. Undoubtedly, the strong association of demand with income is reflecting both motivational factors associated with different student home and school environments as well as cost and ability-to-pay considerations. It is not possible to isolate the effects of the latter factors in this local analysis, since students in the Boston SMSA *ex ante* face an identical schedule of prices at public colleges in Massachusetts and elsewhere. Similarly, these students face a common geographically centered labor market so that asymmetric reactions to the opportunity costs of higher education cannot be effectively discriminated at this disaggregated level. On the other hand, a cross-section regression analysis over high schools in a single urban area does allow one to isolate

-
- 13 The use of a mean sum of squares term rather than simply the mean squared in this nonlinear equation follows directly from our concept that the basic structural equation exists at the individual student level. If a quadratic relation is formulated at the individual student level and then one sums and averages over particular groupings (i.e., schools) before doing the regression, it is easily demonstrated that this leads to relation involving a mean sum of squares term. The grouping technique is a necessary data processing approach in this situation rather than the fundamental unit of observation, and we therefore feel this approach is appropriate.
- 14 The relatively large negative intercept term indicates that there may exist an interval of increasing returns for very low income levels. That is, extrapolation backward using the estimated regression equation (7) yields negative values at low but positive income levels. Because negative values are impossible, this would tend to indicate a range of increasing returns around the origin. However, projections beyond the range of the data are always hazardous.

some of the environmental factors potentially underlying the structural relation that would be obscured at the more aggregate level.

In particular, two empirical proxy variables dealing with the students' home and school environment can be introduced into the current analysis: C_j = current expenditures per pupil in average daily attendance in school j (measured in \$100/pupil); N_j = fraction of nonwhite students enrolled in school j .¹⁵

In order to gain some further insights into the relative role of these different factors in influencing the desire to seek college, we added these two variables to regression equation (7). In the structural form estimated, diminishing returns in the school expenditure per pupil variable was allowed for by including a quadratic term in this variable,¹⁶ while the mean sum of squares for the family income variable was also retained. We obtained the following equation:

$$(8) S_j = -.89 + .160\bar{I}_j + .093C - .42N - .0040\left[\sum_{i=1}^{n_j} (I_{ij}^2/n_j)\right] - .0021C_j^2 \\ \quad \quad \quad (3.30)** \quad (1.94)* \quad (1.82)*(1.33) \quad \quad \quad (1.53) \\ j = 1.56 \quad \quad \quad \bar{R}^2 = 0.58$$

The t -values are in parentheses.

This revised formulation indicates that the first order terms for all three factors have a statistically significant effect (using a 5 percent confidence interval). Nevertheless, the coefficient associated with the mean family income term is still the most dominant factor from the standpoint of statistical significance and elasticity response. The magnitude of the estimated coefficient for this variable declines only slightly (.18 to .16) and the adjusted \bar{R}^2 coefficient rises only moderately (0.54 to 0.58) in comparison to estimates previously observed for equation (7).

15 Expenditures per pupil are included as a proxy variable for the quantity and quality of secondary school inputs received by the students in each school. While this is far from a perfect index of the effect of secondary school environment on a student's motivation and training for college, the supply of many relevant inputs will be correlated with the level of money expenditures. In addition, it may be noted that the schools in the Boston SMSA exhibit considerable variation with regard to current expenditures per pupil. (The coefficient of variation for this variable is 0.25 and the range is from \$625 to 1985.7 with a mean value of 846.) Therefore, this single index may at least capture the gross difference in the supply of these inputs relevant to future college attendance. The other variable, the percentage of students in each school who are nonwhite, is included to capture special factors in the home and school environment affecting the demand for college by minority groups.

16 The second order term of the expenditure per pupil variable may be simply squared in this regression because this variable is ostensibly constant for all students in a particular school. That is, the mean sum of squared term is identical to the mean squared if there is no variation in this variable over the particular group for which it is averaged.

These results therefore indicate that only a moderate portion of the income relatedness of the demand for college can be specifically interpreted reflecting these environmental factors measured in this way. This is not to suggest that these factors have a negligible or insignificant effect. As indicated above, all three variables do indeed have a statistically significant effect. The importance of the income variable is relative in nature.¹⁷ In considering the effect of these other variables, the significantly positive impact of the school expenditure variable (elasticity coefficient of 0.7 at the sample means) is especially interesting. This suggests that expending more funds generally in an income-compensatory instead of the income-reinforcing pattern that currently exists would have some favorable effects on socio-economic mobility. Of course this leaves open the question on what particular policy measures such a reallocation of funds might optimally concentrate.¹⁸ In accordance with a priori expectation, the negative estimate for the quadratic term of the expenditure per pupil variable indicates diminishing returns hold in the effect of this variable on college aspirations. This is also true for the income variable, as the mean sum of squares term is negative. Neither of these higher order terms is statistically significant (however, the *t*-values are significant at the 10 percent level) and they exert a significant impact on the dependent variable only for values at the high end of the spectrum.

The percent nonwhite variable has a negative coefficient which is significant at the 5 percent level. In accordance with the findings of other studies, this result suggests that there are factors such as overt discrimination that make it particularly difficult for certain groups to achieve socio-economic mobility through higher education.¹⁹ On the other hand, unless

17 Well-known econometric problems could be affecting these results. Multicollinearity, while not a severe problem here, could certainly obscure the relative impact of each of these variables. Moreover, variables providing a better index of these environmental influences could also cause these results to change.

18 Often public policy proposals suggest that more funds should be channeled to a particular sector of the economy, such as education, but do not present criteria according to which the funds should be allocated among the various activities of that sector. For example, in the education sector funds may be spent for better teachers, more classrooms, supplies, transportation, etc. Allocation of government funds based upon efficiency criteria requires a refined specification of the inputs and outputs of the educational process and the purchase of those inputs which yield the largest return in student achievement. An analysis of the determinants of the scholastic achievement of high school seniors in the Boston SMSA has identified those factors which yield the highest return in scholastic achievement. These results are reported in A. Corazzini et al. [5].

19 See, for example, G. S. Becker [1, 2].

TABLE 2
PERCENTAGE OF STUDENTS TAKING SAT AND
PERFORMANCE ON VERBAL SECTION OF SAT, BY INCOME QUARTILE

| | Percent Taking SAT by Dec. 1969 | Percent of Those Taking Test Scoring | | |
|----------------------------|------------------------------------|--------------------------------------|---------|---------|
| | | 200–400 | 400–600 | 600–800 |
| Lowest income quartile | 44.9 | 30.0 | 63.1 | 6.8 |
| Second income quartile | 46.5 | 25.2 | 65.8 | 9.0 |
| Third income quartile | 57.6 | 22.1 | 62.2 | 15.8 |
| Highest income quartile | 70.5 | 15.1 | 60.8 | 24.1 |

multicollinearity is obscuring matters unduly, the relative dominance of the income variable in equation (8) indicates that it is lower income itself and the factors generally associated with lower income that are more influential in keeping this disadvantaged group from seeking higher education.

The above results indicate the strong income-relatedness of the decision to seek college enrollment in the Boston SMSA. This is also brought out in Table 2, where our 10 percent sample of 3,200 students are classified according to income quartile (using our constructed index of family income). The first column of Table 2 indicates the percentage of students in each income quartile taking the SAT through December of their senior year. These range from 44.9 percent in the lowest income quartile to 70.5 percent in the highest income quartile.

Given the strong structural relationship between income and the decision to seek college, the next question of analytical interest is the actual SAT performance by students of different income groups. In Table 2, a breakdown of scores on the verbal portion of the SAT is presented for the students in each income quartile taking the exam. As one can readily see from these figures, performance on the SAT is also strongly income-related. The two lowest income quartiles, which have far fewer students attempting the exam, nevertheless exhibit much poorer performances. In the lowest income quartile, 30 percent of the students taking the exam fall below 400, whereas only 6.8 percent score above 600. The second lowest income quartile does only marginally better.

The significance of these results can best be appreciated by realizing that well over 95 percent of enrolling freshmen at the state university and the various four-year state colleges in Massachusetts score over 400 on the verbal portion of the SAT. Furthermore, Massachusetts state scholarships are normally awarded only to candidates scoring at least an average of 650

on both the verbal and math portions of the SAT. Therefore, the group which is in most need of financial support, the lowest two income quartiles, is in the poorest position to realize it—at least in the form of subsidized education at state-supported schools and outright scholarship grants.

The breakdown of SAT performance presented in Table 2 indicates that the majority of those students seeking college admission who will be frustrated by aptitude constraints will be lower-income students. However, many of these students may succeed in getting into junior colleges where admission requirements are less stringent. While only 9.1 percent of the students desiring college from the lowest income quartile expressed an interest in attending a junior college, many of them may be “bumped” down into a junior college program, as well as out of the college-going stream. While this “bumping-down” phenomenon is obviously preferable to complete frustration of college plans, it hardly can be viewed as an optimal long-run solution.

V. SOME POLICY IMPLICATIONS

The local analysis reinforces the regression results of the national cross-sectional analysis in two important respects. First, it is quite clear that family income is important in determining who enters the market for a college degree. Beyond environmental effects, family income serves, at least in part, as an effective bypass to the imperfections of the capital market with the end result that students from low-income families—even when qualified—are often unable to attend college. Expanded scholarship and loan programs would seem to be necessary. Second, high school students from the lowest socioeconomic quartile do indeed face a severe admission standard constraint. Therefore, traditional scholarship and loan programs, taken by themselves, are inappropriate policy tools if the objective is to induce large numbers of students from such backgrounds to enter four-year colleges. It is possible, however, that the existence of such programs on a very large scale might serve to grind down this barrier, insofar as academic performance was directly related to a promise of complete funding. Beyond this, the national results indicate that the opportunity costs of college attendance are not given sufficient attention by those planning higher education financing. These costs are crucial to those in the lowest socioeconomic quartile and are not taken into account in so-called free tuition plans.

The current proposals to establish a voucher system in higher education are also inappropriate in this regard. Voucher systems provide the partial economic means by which students overcome a financial constraint, however they abstract from the problems of admission standards at the various

schools and the opportunity costs of attendance.

Unless one is willing to argue that the SAT is a completely accurate measure of ability, the distribution of the college population between prestige colleges and two-year junior colleges will be in large part a simple reflection of socioeconomic class. It is quite clear that SAT scores are not designed to be an index of native ability, but rather are a measure of performance. The very fact that both environmental and educational influences upon these scores can be identified and in some way measured indicates that the scores are not pure measures of inherent aptitude or intelligence.²⁰ This is not to deny that measured performance on such tests may successfully reflect certain skills needed to meet conventional academic demands in institutions of higher education, or that such scores may even predict expected achievement.

These results are all the more disturbing in the light of recent evidence regarding the higher education expectations of the college-age population. In 1959 and again in 1965, the U.S. Bureau of the Census in cooperation with the U.S. Office of Education conducted national surveys of high school seniors concerning their aspirations for post-secondary education. The two surveys showed that the aspirations of students from low-income families had changed dramatically. In 1959, 23 percent of high school students from families with incomes of less than \$3,000 hoped to attend college, but by 1965, 46 percent of all students from families in an equivalent real income range held college expectations. This doubling in the rate of hoped-for college attendance by the low-income group contrasts sharply with a 6 percent increase in the expectations of students from high-income families. Overall, 60 percent of seniors hoped to attend college in 1959, and by 1965, 70 percent held that expectation. This increase reflects this rapid change among students from low-income backgrounds. The Boston SMSA data also reflect this nationwide shift. Sixty-nine percent of all seniors and a full 60 percent of all those in the lowest income quartile held college expectations.

Even if we assume for the moment that these rates of hoped-for college attendance can be taken as a measure of the potential demand for higher education, it is not possible to evaluate current public policy in the light of the subsequent aggregate rationing that occurs. As we have pointed out, aspirations are simply that, and some rationing will occur. It is, however, the disproportionate rationing of low-income students from places in four-year colleges that gives us cause for concern. Assuming the SAT is an efficient measure of potential college success and one does not wish to alter radically the present college program, means ought to be found to qualify large numbers of low-income students for four-year colleges and to finance

20 See, for example, S. Bowles and H. M. Levin [3].

their college careers. The resource costs of many of the solutions which come to mind (that is, an extra year of schooling for such students, special tutoring, etc.) would be considerable. Nevertheless, the social costs of not adopting such programs may be even greater.

APPENDIX: SOURCES OF NATIONAL CROSS-SECTION DATA

Socioeconomic Quartiles. Nine items were used to construct the Project Talent Socioeconomic Environment Index. These items were selected to measure aspects of the environment closely related to student ability and to control for home socioeconomic conditions in studies of school learning. The nine items include value of home, family income, books in home, appliances, TV, own room, father's occupation, father's education, and mother's education. For a detailed description of the index, see *Project Talent: One-Year Follow-Up Studies* (Pittsburgh: School of Education, University of Pittsburgh, 1966).

Enrollment Proportions. Proportions of 10th grade high school students who enrolled in college in 1963 were obtained from *Project Talent: One-Year Follow-Up Surveys* of 1960 high school students.

Tuition Costs. Tuition costs for each type of institution are state averages obtained from U.S. Office of Education, National Center for Educational Statistics, unpublished data from Higher Education General Information Survey, 1963-64.

Performance on Achievement Tests. Ability is a composite variable determined by several scores (verbal and mathematical achievement plus general information) derived from *Project Talent Survey* of 1960 high school students.

Paternal Education. Father's education attainment level was obtained from *Project Talent: One Year Follow-Up Surveys* of 1960 high school students. Seven levels of educational attainment for the student's father were included on the survey; these ranged from less than 8 years to 5 or more years of college. These years of school completed are associated with mean income in the U.S. Bureau of the Census, *Current Population Reports*, Series P-60, No. 56, "Annual Mean Income, Lifetime Income, and Educational Attainment of Men in the U.S., for Selected Years, 1956-1966" (Washington: U.S. Government Printing Office, 1968), Table 1, p. 27. For 1963, the mean income associated with the levels of educational attainment are presented below (in current dollars):

| <i>Men, 35-44 Years Old</i> | <i>1963 Income</i> |
|-------------------------------|--------------------|
| Elementary: less than 8 years | \$ 3,838 |
| 8 years | 5,169 |
| High school: 1 to 3 years | 5,906 |
| 4 years | 7,122 |
| College: 1 to 3 years | 8,144 |
| 4 years | 10,525 |
| 5 years or more | 11,020 |

Unemployment Rates. Average annual unemployment rates by state are available in U.S. Department of Labor, *Manpower Report of the President*

(Washington: U.S. Government Printing Office, 1968), Table E-4, p. 282.

Wage Rates. Average hourly earnings of production workers on manufacturing payrolls in 1963 were used to represent the wage rate pertinent to the high school student's decision to attend college. These data were drawn from Bureau of Labor Statistics, *Handbook of Labor Statistics, 1968* (Washington: U.S. Government Printing Office, 1968), Table 85, p. 173.

REFERENCES

1. Becker, G. S. *Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
2. ———. *Human Capital*. New York: Columbia University Press, 1964.
3. Bowles, S., and Levin, H. M. "The Determinants of Scholastic Achievement—An Appraisal of Some Recent Evidence." *Journal of Human Resources* 3 (1968): 3–24.
4. Campbell, R., and Siegel, B. N. "The Demand for Higher Education in the United States, 1919–1964." *American Economic Review* 57 (1967): 482–94.
5. Corazzini, A. J., et al. *Higher Education in the Boston Metropolitan Area: A Study of the Potential and Realized Demand for Higher Education in the Boston SMSA*. Boston: Board of Higher Education, 1969.
6. Hansen, W. L. "Income Distribution Effects of Higher Education." *American Economic Review* 60 (1970): 335–41.
7. Galper, H., and Dunn, R. M., Jr. "A Short-Run Demand Function for Higher Education in the United States." *Journal of Political Economy* 77 (1969): 765–77.
8. Schultz, T. S. "Investment in Human Capital." *American Economic Review* 51 (1961): 1–17.
9. U.S. Office of Education. *Digest of Educational Statistics*. Washington: U.S. Government Printing Office, 1969.