SCHOLASTIC APTITUDE TEST SCORES AND THE ECONOMIC RETURNS TO COLLEGE EDUCATION

by

Charles Noah Steele

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 1990

APPROVAL

of a thesis submitted by

Charles Noah Steele

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Date Chairperson, Graduate Committee

Approved for the Major Department

Date Head, Major Department

Approved for the College of Graduate Studies

Date Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor or, in his absence, by the Dean of Libraries when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature______________________________

Date______________________________

ACKNOWLEDGEMENTS

My thanks go to my graduate advisor, Professor Douglas J. Young. Without his help this thesis would not have been possible. I am also grateful to my two other graduate committee members, Professors Ronald Johnson and Vincent Smith, for the valuable advice and suggestions I received from them. Finally, I would like to express my gratitude to the other faculty members and graduate students who aided me in this project. Thanks.

TABLE OF CONTENTS

TABLE OF CONTENTS -- Continued

LIST OF TABLES

ABSTRACT

Student scores on standardized achievement tests fell during the 1960's, raising questions about the quality of education in America. The decline was especially pronounced on the Scholastic Aptitude Test (SAT), the most widely used college entrance examination. Much of this decline remains unexplained.

This study examines the hypothesis that declining returns to college education played a role in the score decline. Specifically, it is hypothesized that declining returns to education reduce the incentive for students to invest in the college skills which the SAT attempts to measure. This study proposes a general model in which SAT participation rates and SAT scores are determined in part by the relative wage differential between an average college graduate and a high school graduate. Several empirical specifications of this model are then estimated, using aggregate time series data for the years 1967 to 1987. The methods of three stage least squares and seemingly unrelated regressions are used.

The results of the regressions suggest that fluctuations in the relative economic return to a college education may have accounted for between 11 percent and 24 percent of the decline in average SAT scores. When these effects are considered in conjunction with those of a demographic variable, approximately 40 percent of the decline is explained, perhaps suggesting that fears of declining educational quality are somewhat exaggerated.

CHAPTER 1

INTRODUCTION

Problem: The Unexplained Nature of SAT Score Fluctuations

The Scholastic Aptitude Test (SAT) is the most widely used standardized examination for college admission in the United States. Approximately one and one half million examinees, largely high school students, participate in this testing program each year. The SAT is widely recognized as a measure of student achievement, and is frequently regarded -- perhaps not always properly -- as an indicator of how well American schools are performing.

In 1963, average SAT scores began a precipitous decline which continued until 1979. This decline was not simply an artifact of changes in the SAT itself, nor of the group of students self-selecting to take the exam: throughout the sixties and seventies scores on a number of other standardized achievement tests also showed substantial declines. Since 1979, SAT scores, and other achievement scores, have recovered somewhat, although not to pre-decline levels.

The achievement decline, and especially the SAT score decline, provoked great alarm and discussion, and was important in the development of a public debate about educational policy which continues today. Parents, educators, and policy makers spoke worriedly of falling educational standards, failing educational systems, and social changes detrimental to learning, and often prescribed a multitude of corrective measures. Researchers enquiring into causes of the decline proposed a wide range of explanations. (For example, a survey undertaken in the mid-seventies found some 79 separate hypotheses which had been offered to that date to account for the SAT score decline. Wharton, 1977.) Without question the fall and subsequent rise of SAT averages has been and continues to be of great importance in educational policy debates.

It is noteworthy that despite the attention which has been devoted to SAT fluctuations, the fall and partial recovery of scores has not been adequately explained. Of those hypothesized causes which can in some way be tested, many appear to have had only a small effect on test score trends, while many more seem to have had no impact at all. Indeed, throughout the literature devoted to testing issues there are frequent references to the unsolved nature of the decline. The question is still very much open as to what factors caused SAT scores to fall, and why the trend later reversed.

The hypotheses proposed to explain the score decline have come from many fields, and include such disparate phenomena as changes in the conduct of classes, changes in popular attitudes towards learning, various demographic effects associated with larger family size and the baby boom, increased television viewing, and environmental pollution. But despite this wide range of proffered explanations, little attention has been devoted to possible explanations from economics. This is especially interesting, and perhaps somewhat surprising, given the extensive work which has been done in the economics of education. This work has clearly shown that education can often be thought of as an investment in human capital, an investment which is readily studied with standard economic analysis of equating returns and costs at the margin.

Early work by Becker (1964) looked at investment in human capital by individuals from the standpoint of income maximization. Becker's approach was quite successful in explaining a wide variety of forms of investment in education and training, and his model provided the basis for much of the subsequent work in the economics of educational investment. In his survey of the work of Becker and those who followed, Richard Freeman (1986) observes, "The investment model of the decision to pursue education has been studied extensively with generally favorable results for the key behavioral assumption: that individual decisions respond significantly to meet incentives."

If individuals significantly alter their investment choices in response to changing economic incentives, it makes sense to investigate the hypothesis that changing economic incentives played a role in the SAT decline. It is the aim of this study to do just that. As will be seen, during the period of the SAT decline there was also a substantial decline in the earnings of a worker with a college education relative to those of a worker with only a high school education. If individual investment in college education decreased in response to this decline, so too might have individual investment in the "college skills" which the SAT attempts to measure. And to the degree that the decline in the return to college is a temporary phenomenon, so might be any SAT decline induced by it, which would argue against the need for some strong policy response. While such an hypothesis from education might not fully explain the general decline in student achievement (especially in the lower grades) it could potentially go far in accounting for SAT score fluctuations.

The Approach of This Study

This study will begin by examining the work which has been done to date on SAT score trends. Approaches which have proved productive will be reviewed, as will some of the problems and inadequacies of previous research. The economic literature suggesting the possibility of an economic factor in SAT fluctuations will also be briefly discussed.

In Chapter 3 a formal analytical model relating student achievement with returns to education will be presented. Implications for SAT score trends will be examined.

The remaining chapters will examine the consistency of the proposed explanation with the empirical data. Chapter 4 will describe the data which will be employed in this study. Factors which are hypothesized to be important in explaining scores will be discussed, as will be the generation of the actual explanatory variables which will be used in econometric estimation. Trends in the data over the time period covered by the study will be briefly examined, and a formal model for estimation presented. Chapter 5 will describe the results of the estimation procedure, and Chapter 6 will discuss what conclusions might be drawn from the results.

Some Facts About the SAT

The SAT is the most widely used college entrance examination in the United States. It is given to students across the nation several times each year under the auspices of the Educational Testing Service (ETS), which designs the exam and oversees its administration.

Student participation in the Scholastic Aptitude Testing program is voluntary. Not all universities require that prospective entrants take the SAT; some give students the choice of taking the SAT, the ACT (American College Testing Program), or some other examination, while other universities require no test at all. The SAT is typically used as a predictor of a student's future performance in college. Universities using the SAT as an application credential usually weigh the scores in conjunction with other information concerning a student's academic ability, such as grade point average.

The Scholastic Aptitude Test consists of two parts, verbal and mathematics, and is sometimes thought of as two separate tests. On each section there is a maximum of 800 points possible, and a minimum of 200. A score of 600 or above on a section is generally considered high.

The scores are scaled in such a way as to keep the scale constant from year to year. Thus, scoring 500 on an exam in any one year is equivalent to scoring 500 in any other year. The scale is calibrated for each individual test by means of common questions scattered through consecutive exams. One hypothesis offered by researchers to account for the SAT decline was that of scale drift. However, inquiries into this possibility found that if anything the SAT scale had drifted in such a way as to make achieving any given score slightly easier. (Congressional Budget Office, 1987). If so, the score decline may have understated any decline in student ability.

Because participation in the testing program is voluntary, the SAT-testing population does not represent a random sample from the student population. SAT-takers presumably tend to be drawn from the higher achieving students, who are more likely to be preparing for college, especially the more exclusive colleges which are more likely to require the SAT. If the proportion of a cohort of students taking the exam increases, it is generally thought that more less-able students will be entering the SAT-taking population, and that the average score will fall even though there may have been no decline in student achievement. This selectivity effect must be taken into account in modeling and estimation, and, incidentally, is an important reason why using SAT scores as a gauge of the performance of schools may lead to erroneous conclusions.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

This literature review will examine some of the fundamental issues involved in understanding SAT score fluctuations. The survey offered here first considers two recent publications which broadly summarize what has been learned about achievement score fluctuations in the last two decades. Next, the findings of a number of important articles dealing with various aspects of the SAT decline are briefly examined. Finally a short review at the literature of the economics of education lends support to the idea that the scores may have fluctuated due to economic factors.

The CBO Studies

Two surveys of the fluctuations in test scores and hypotheses proffered to explain them are the studies by the Congressional Budget Office "Trends in Educational Achievement" (CBO, 1986) and "Educational Achievement: Explanations and Implications of Recent Trends." (CBO, 1987) The 1986 study documents the trends in a variety of achievement test scores since the 1960's. The start of the SAT decline in 1963 initiated a period of generally falling achievement test scores. In 1966 declines began in the American College Testing Program (ACT), the Iowa Test of Basic Skills (ITBS, given to grades 5 and 8), and the Iowa Test of Educational Development (ITED, grade 12). The declines in the Iowa tests are noteworthy as the test-taking populations for these are not self-selected; all students in each grade take the tests.

Among the general findings are several other important points. First, it is clear that the decline in standardized achievement test scores was not limited to college entrance examinations such as the SAT. However, the general decline was most pronounced in the upper grade levels, and appeared to occur primarily in "higher order skills", i.e. reasoning and problem-solving; the skills which the SAT attempts to measure.

Second, the study found limited evidence to suggest a period effect in the decline and subsequent rise in test scores. The term "period effect" refers to a phenomenon which affects all grades and birth cohorts simultaneously. A period effect would entail a change in some factor which could be observed by or otherwise influence all birth cohorts at the time when the change occurs. A change in the economic incentives to acquire intellectual capital might be thought of as a period effect, if the change is exogenous to every cohort. It will be seen below, however, that there is reason to believe that relative cohort size has a strong effect on returns to education.

On the other hand, the study found evidence from the ITBS, the ITED, and the SAT which points to a substantial cohort effect in the fluctuations in student achievement. For example, the reversal of the decline appears to have been initiated by the 1962-1963 birth cohorts, beginning in the lower grades as they entered school, and continuing into the higher grades as they moved up through the educational system. This would seem to imply that some change specific to the 1962-1963 birth cohort occurred, and persisted, or was repeated, for following cohorts.

The 1987 CBO study analyzed a number of the most likely explanations of the trends in test scores. Although a large number of hypotheses have been offered, many are difficult or impossible to test, and of those which have been to some extent tested, many appear to have had little or no impact on score trends. The study divided the more plausible explanations among them into three basic categories: educational factors, selection factors (changes in the population of students being tested) and general social and cultural trends.

Educational factors which the CBO authors examined as possible causes for the fall and rise include changes in teachers' levels of experience, textbook changes, federally funded educational programs for students from low income families, changes in course content, and the amount of homework. Empirical evidence for the effect of teacher experience on student achievement has been contradictory and weak. Federal program changes did not coincide with the fall in scores and cannot account for it, but may have had some affect in recent SAT score gains by minority groups. Evidence for changes in course content is largely anecdotal and inconclusive.

Social and cultural changes which were considered include ethnic composition, family size, drug use, and environmental lead. The CBO authors report that changes in ethnic composition of the student population may have been responsible for one tenth to one-fifth of the decline, and may also hindered the recovery. Changing patterns of alcohol and drug use may have been a factor in both the fall and rise of scores, especially for the higher grades, and reduced levels of lead in the environment may have aided the upswing. Evidence for these was found to be weak. Family size may also have contributed to the fluctuations, with some estimates of the effect on the score decline as high as twenty-five percent.

Self-selection almost certainly was a factor in the changing scores on college entrance examinations, but would have had no effect on scores for other tests. Self-selection occurs only for such exams as the SAT and ACT, but the score decline was evident for such tests as the ITBS and the ITED as well.

The study also listed a number of proposed contributing factors which appear to be ruled out on the basis of evidence, or for which there is simply insufficient evidence to suggest any conclusion. Included in the factors for which evidence suggests little or no effect are several measures of teacher ability, changes in state graduation requirements, increasing numbers of single parent households and working mothers, and increased television viewing. Among suggested factors for which evidence was not available are teacher attitudes and morale, local graduation requirements, grade inflation, and student attitude and motivation.

The conclusion of the study is that no single cause accounts for the drop and subsequent rise in achievement scores. Rather, the phenomenon seems to be attributable to a variety of factors working in concert. Furthermore, much of the fluctuation remains unexplained.

The 1987 study also makes an important point concerning the criteria by which proposed explanations are evaluated. Factors being suggested must be both cross-sectionally and temporally consistent, that is, they must be both demonstrably related to student achievement at any given time, and they must correlate in time with the trends in achievement scores.

Specific Studies

Early work done on the question of the test score changes suggested a very important role for demographics; specifically, for changes in family size as a result of the postwar baby boom. Psychologist R.B. Zajonc (1976) argued that changes in family configuration could account for a large portion (but not all) of the decline in SAT and other scores. Noting the likely environmental effects of family size on the development of children's intelligence -- increased intellectual performance with decreasing family size, earlier born children generally outperforming younger siblings (the "confluence model"; Zajonc and Markus, 1975), Zajonc pointed out the close correlation between changing average birth order in American families and the SAT decline.

However, later studies using cross-sectional data suggested that Zajonc's findings overstated the effects of family size. In a study done for ETS on the Zajonc hypothesis, Hunter Breland (1977) was unable to conclude that changes in family configuration contributed strongly to the score decline, and that without further evidence the effect could be assumed to be small. In later study, Zajonc concluded that the effect of family size accounted for less than ten percent of the drop in scores (Zajonc and Bargh, 1980). Zajonc also observed that from the other hypotheses offered it was difficult to find one which would account for both the decline and the (then anticipated) rise of test scores and he suggested that perhaps two unrelated (and unknown) factors were at work.

In a more recent study of the relationship between family size and intellectual performance Eric Hanushek (1987) concluded that there is little advantage to being earlier in birth order, holding family size constant. However, he also concluded that there does exist a definite advantage to children in smaller sized families, due to the increased proportion of parental attention available to each child. He speculated that family size changes might account for more than fifty percent of the change in Iowa Test scores, and as much as twenty percent of the change in SAT averages. He also found the effects of increased numbers of working mothers and single parent families to be unimportant.

William Schrader (1976) examined changing patterns of college attendance and SAT-taking as possible factors in the SAT decline, utilizing data on the number of SAT score reports sent to various types of institutions of higher education in 1960-61, 1966-67, and 1973-74. He suggested that a trend of increasing percentages of high school graduates attending college coupled with the increasing acceptance by colleges of either the SAT or the ACT was decreasing the proportion of SAT-takers who were higher ability students. He also noted a substantial increase in score reports to two-year colleges, specialized colleges, nursing schools, and other institutions which might be thought to attract less-able students.

Utilizing data from both the SAT verbal and quantitative test scores, Jackson (1976) uncovered several points of interest. Prior to 1971 the numbers of high-scoring students (i.e. over 600) on each exam section did not decline, even though the percentages of students achieving high scores did fall. This would be consistent with the hypothesis that the decline in average scores was the result of more less-able students taking the examination, i.e. a "selection" or "proportion" effect. Yet beginning in 1970-71 the absolute number of high-scoring students did decline, while at the same time the percentage of SAT-takers reporting high levels of school achievement increased. Jackson was not able to ascertain to what degree changes in test-taking may have contributed to the score decline, but concluded that such changes could not have been a sole cause.

William Turnbull (1985) attempted to formulate a comprehensive explanation of the test score decline. He argued that schools have increasingly retained and graduated less-able students, especially since World War II, largely as the result of increasing public expectations as to the proper amount of schooling. In Turnbull's view, the first stage of the score drop was simply the result of increasing numbers of marginally qualified prospective college entrants taking the examination. Turnbull also argued that high schools began to "dumb down" programs in order to retain less-able students, and that this further contributed to the decline, and especially to the second stage of the decline, the drop in the scores of high-achieving students. Turnbull's explanation appears to be logical and not inconsistent with the data, but much of his evidence (decreasing quality of texts and classes, changes in student motivation, societal pressures, etc.) is sketchy or anecdotal. Turnbull also criticized the idea that there may have been two unrelated declines (the initial decline in the SAT average followed by the decline in high scores), arguing that the observed decline was a single, continuous phenomenon.

Mark Dynarski (1987) used pooled cross-sectional time series data at the state level to examine the effects of various factors on SAT scores, and found a strong relationship between average scores and participation rates. His estimates indicated that a one percent increase in student participation corresponds to an approximately 1.8 point decline in the average score. Such a finding is strong evidence for the hypothesis that increased participation implies more less-able students in the test-taking population, giving lower average scores without necessarily indicating declining levels of student achievement.

Bishop (1989) recently examined a possible relationship between intellectual achievement as measured by standardized examinations and worker productivity, hypothesizing that diminishing levels of skills among workers entering the labor force might be responsible for productivity growth declines in the United States. He concluded that the impact of declining educational achievement on wages and productivity was much more significant than expected, and estimated the cost in forgone GNP to be 86 billion dollars in 1987.

Bishop did not address possible causes of the score decline, and suggested that it remained a mystery. He noted that the declines in productivity growth and test scores occurred simultaneously, and speculated that the cause might lie outside education in such things as changing attitudes toward school and work. It should be noted that Bishop's approach is quite different from that of this study. He viewed lower levels of achievement as a possible cause of productivity decline, while the present study investigates declining returns to education as a possible cause of lower achievement levels.

Several points emerge from this discussion. First, the question of why SAT scores fluctuated has not been resolved. Although a number of hypotheses have been suggested, they are frequently inconsistent with the data or could not have accounted for both the fall and rise of test scores. More than one researcher has suggested a sort of "missing link", an unknown variable which could have been responsible for the fluctuations. Second, it appears likely that there was no single cause acting alone, but rather that several factors operating together were responsible for the fall and rise. Breland (1977) speculated on the possibility of an unknown intervening variable through which suspected causes might be linked to the score decline. And also, it appears that selectivity has played an important role.

The Economics of Education Literature

It is interesting to note that none of the explanations for the SAT decline offered above are based on economic theory, even though student motivation is frequently cited as a likely (and untestable!) contributing factor to the fluctuations. This is despite a large body of work by economists on education as an investment in human capital. For example, the work of Gary Becker (1975) has shown that college graduates earn more than high school graduates in large part because of their additional education, and that such returns on education are a strong motive for obtaining education. If student motivation is an unexamined potential factor affecting SAT scores, and a primary reason for obtaining education is the economic benefits accruing to such an investment in human capital, then an examination of relative costs and benefits might shed light on this aspect of the question.

Finis Welch (1979) investigated the relation between relative cohort size and earnings, and found that larger cohorts entering the labor market tend to face lower wage rates, due to the increased supply of workers. Especially interesting is his finding that the decline in wages is greater for more highly educated workers, and that the effects do persist somewhat throughout the career. These results may suggest links between cohort size, returns to education, and SAT scores. As will be seen, the lowest points of the SAT decline corresponded closely to the years in which the peak baby boom cohorts passed through high school.

Similarly, Stapleton and Young (1988) argued that baby boom cohorts can minimize the detrimental effect of crowding on their earnings through optimal educational and career choices. Their study corroborated Welch's findings that entering workers are poorer substitutes for experienced workers in occupations requiring higher levels of education, and that therefore large cohorts entering the job market will depress wages for new entrants more in such occupations. Stapleton and Young then showed that boom cohorts can overcome such declines by investing less heavily in education, while pre- and post-boom cohorts can benefit by obtaining more education.

These ideas suggest a rough model of how returns to education might influence SAT scores. A decline in the relative return to education, perhaps brought about by increased cohort size, would induce less investment in human capital by students. Especially, there would likely be less investment in those sorts of skills specific to college and college careers. These lower levels of achievement might well be reflected in SAT scores. Such a scenario would not indicate a problem within education at all, but rather the rational responses of individuals to changing market conditions. Investigation of these ideas will require a formal analytic model.

CHAPTER 3

A FORMAL MODEL OF INVESTMENT IN COLLEGE SKILLS

Introduction

In this chapter a formal model of the decision-making process for an individual high school student regarding post-high school education choices is developed. In this model the economic return to college education helps to determine the level of intellectual capital or skills that the individual will acquire. It is assumed that this capital is acquired by the student while in high school, and consists of those skills specific to successful pursuit of a college degree. It is further assumed that these are the skills which the SAT measures. From the individual model an aggregate model for the entire cohort is constructed. The model is analyzed for implications concerning the relation between SAT scores and economic incentives to acquire college-type skills.

Initial Considerations

A formal economic model of college choice behavior should capture the response of the individual's decisions to changing economic incentives. In general, individual behavior is better modeled as utility maximization rather than income maximization with fixed labor supply. For example, if the wage rate should increase, an individual might choose to remain at the same level of income, substituting more leisure for less work (or study).

However, there are difficulties in using utility maximization which in some cases may make an income maximization model appear more desirable. Utility, based upon the subjective preferences of the individual, cannot be observed. Accurate data on time devoted to leisure by students is difficult or impossible to obtain. Income is thought to be the driving force in individuals' educational investment decisions; certainly previous work using an income approach to model educational investment has proved very successful (e.g. Becker, 1975). The present study, then, will utilize income maximization in modeling individual college choice and test-taking behavior, primarily for reasons of tractability.

Modeling Individual Decisions

The model of college choice presented here is one in which the individual compares his lifetime earnings opportunities with and without a college degree. The student will select the track -- college or high school -- which offers him the greater return. The form of the model is as follows:

Go to college if and only if:

$$Y_i^c \geq Y_i^h \tag{3.1}$$

where

$Y_i{}^c$ = potential earnings, individual i, college track, and

$Y_i{}^h$ = potential earnings, individual i, high school track.

These earnings potentials can be expressed in greater detail. A more complete specification of the high school track earnings is given by:

$$Y_i^h = T^h w_h + T^s w_h \tag{3.2}$$

where

w_h = dollars earned per unit time if an individual does not attend college, i.e. the high school track wage; also, the wage while in high school;

T^h = total time an individual works at high school wage after completing high school, if high school track;

T^s = time an individual allocates between study and work while in high school.

The high school track earnings are thus simply the product of the wage rate and the time devoted to work. This time includes time devoted to work while in high school as well as after graduation.[1]

College track earnings are given by:

$$Y_i^c = w_c k_i^* T^c + (T^s - \theta_i^*) w_h - C \qquad (3.3)$$

where

w_c = dollars earned per unit time of intellectual ("college") capital; i.e. the base college "wage";

k_i^* = the optimal level of college capital individual i will acquire if he chooses the college track;

T^c = total time an individual works at college wage, if college track;

θ_i^* = optimal amount of T^s allocated to study for individual i, if college track; $\theta_i^* \leq T^s$, and,

C = direct costs of college (tuition, books, etc.).

Lifetime earnings for the college track are the sum of earnings while in high school at the high school wage, and earnings after college at the college wage. Earnings at the college wage are influenced by the level of intellectual capital acquired while in high school, k_i. The level of college capital, in turn, is determined in part by the amount of time devoted to study. This study time is

---

[1]The wage while in high school could be different from the wage after high school. However, nothing is lost by assuming them to be the same, because the focus is on high school versus college.

subtracted from the time available for work while in high school, a total which is assumed fixed. The amount of this time which an individual devotes to study is θ_i. The remainder of the time is devoted to work at the high school wage.

Although the model focusses on only one aspect of a student's decision -- how much time to devote to study -- it is intended to represent a broader set of decisions, such as whether to take college preparatory type courses. These courses are usually more difficult and require a greater amount of time than other courses, and thus involve essentially the same tradeoff between study time and time that could be devoted to other activities, including market work.

The level of college capital, k_i, can be thought of as depending on two factors which vary across individuals: the amount of time devoted to study, θ_i, and ability, a_i. Ability refers to all inherent attributes of the individual, genetic, environmental, or otherwise, which affect his intellectual development. Thus, ability may be thought of as exogenously given, while the amount of time devoted to study is an endogenous variable over which the individual optimizes. In this way college capital can be seen as a function of both ability and effort devoted to study, or:

$$k_i = f(\theta_i, a_i) \tag{3.4}$$

where the first and second partial derivatives are hypothesized to be as follows:

$f_\theta > 0$

$f_a > 0$

$f_{\theta\theta} < 0$

$f_{\theta a}$ and f_{aa} of undetermined sign.

The positive signs on the first partial derivatives indicate that increasing the time devoted to study and higher ability are each expected to raise the level of intellectual capital acquired, ceteris paribus. The sign on the second partial derivative with respect to θ_i is negative, reflecting the assumption of diminishing marginal returns to study time. Neither the second partial with respect to ability nor the cross partial are given any sign; there is no a priori expectation with regard to either. The second partial derivative with respect to a_i would be negative if there are diminishing marginal returns to ability. However, it is not clear whether intellectual achievement increases at an increasing, constant, or decreasing rate for students of higher ability, and no assumption is made in this model. Similarly, $f_{\theta a}$ is not signed. It is not clear whether the marginal return to study time in terms of achievement is higher or lower for students of higher ability.

For an individual i evaluating the college track alternative, the issue is one of maximization of income over k_i. From (3.3) and (3.4) this can be stated as:

$$\max_{\theta_i} = Y_i^c = w_c f(\theta_i, a_i) T^c + (T^s - \theta_i) w_h - C \qquad (3.5)$$

The first order condition for an interior maximum is:

$$\frac{\partial Y_i^c}{\partial \theta_i} = w_c f_\theta T^c - w_h = 0 \qquad (3.6)$$

The first term on the right hand side of (3.6) is the base college wage rate per unit of intellectual capital, times the increment in college capital derived from increasing the time devoted to study, times the amount of time the capital will be employed. This gives the marginal benefit of an increase in θ_i. The second term is the high school wage rate, i.e. the marginal cost of increasing study time. This equation can be rearranged with the following result:

$$\frac{w_c}{w_h} f_\theta = \frac{1}{T^c} \qquad (3.7)$$

Solving this equation for θ_i^* results in the following statement:

$$\theta_i^* = g(R, a_i, T^c) \qquad (3.8)$$

where R is the ratio of w_c to w_h, i.e. the relative college wage.

It can now be seen that the change in θ_i^* induced by an increase in the relative college wage is positive.

Differentiating (3.7) with respect to θ_i^* and R and rearranging gives the following statement:

$$\frac{\partial\theta^*_i}{\partial R}=-\frac{f_\theta}{Rf_{\theta\theta}} \tag{3.9}$$

From (3.4) it is clear that this is greater than zero.

The effect of a change in ability on θ_i^* is not clear. Differentiating (3.7) with respect to θ_i^* and a_i and rearranging gives the following result:

$$\frac{\partial\theta^*_i}{\partial a_i}=-\frac{f_{\theta a}}{f_{\theta\theta}} \tag{3.10}$$

The sign of (3.10) cannot be determined without further assumptions about the sign of $f_{\theta a}$. If the marginal return to study time in terms of skills acquired is higher for higher ability students, then this statement is positive, and higher ability students will study more than lower ability students, all else equal. Similarly, if the marginal return to increased study diminishes with higher ability, or is zero, higher ability students will study less, or the same amount, as less-able students.

Equation (3.8) gives the optimal amount of study time, θ_i^*, for an individual if the individual plans to go to college. When this value is inserted into the college capital "production function" (3.4), the optimal level of capital for the individual, k_i^*, is determined. Thus, from

(3.1), an individual will attend college if and only if:

$$Y_i^c \equiv w_c f(\theta_i^*, a_i) T^c + (T^s - \theta_i^*) w_h - C \geq w_h T^h + w_h T^s \equiv Y_i^h \qquad (3.11)$$

The influence of the college wage and ability on whether an individual attends college can be examined by differentiation of (3.11). Because these factors do not influence an individual's potential earnings if the high school track is chosen, it is sufficient to consider the effects on income if the college track is chosen.

$$\begin{aligned}\frac{\partial Y_i^c}{\partial w_c} &= f(\theta_i^*, a_i) T^c + w_c f_\theta T^c \frac{\partial \theta_i^*}{\partial R} \frac{1}{w_h} - w_h \frac{\partial \theta_i^*}{\partial R} \frac{1}{w_h} \\ &= k_i^* T^c + (w_c f_\theta T^c - w_c) \frac{\partial \theta_i^*}{\partial R} \frac{1}{w_h} \qquad (3.12) \\ &= k_i^* T^c > 0\end{aligned}$$

The first term on the right hand side of (3.12) is the "direct" effect of an increase in the college wage on income, given some amount of college capital, k_i^*. The second and third terms reflect the fact that an individual's optimal time allocation while in high school is affected by the college wage. However, the condition for optimal time allocation in (3.7) implies that the sum of these two terms is zero. Thus, it is clear that an increase in the college wage increases the potential income of the college track, and thus the attractiveness of attending college.

Similarly, the effect of higher ability on the potential income of the college track is positive. From (3.11) this is given by the following:

$$\frac{\partial Y_i^c}{\partial a_i} = w_c f_a T^c + w_c f_\theta T^c \frac{\partial \theta^*_i}{\partial a_i} - w_h \frac{\partial \theta^*_i}{\partial a_i}$$

$$= w_c f_a T^c > 0 \qquad (3.13)$$

Thus, individuals with higher ability have greater potential college income and are more likely to attend college.

The results of this basic model of individual choice can be summarized as follows. The i^{th} individual will attend college if and only if his potential earnings from the college track exceed those from entering the job market directly after completing high school. The choice variable for the individual should he choose college is the fraction of time devoted to accumulating college skills, and he will allocate this time in such a way as to maximize his total income.

Given a fixed amount of time to be allocated between study and work while in high school, an increase in the monetary return to college capital will lead to increased study and thus more college capital for an individual choosing the college track. An increase in the college wage will also increase the probability that an individual's optimal potential college capital will be above the minimum he would need to make the college track an economically sound choice. Thus a higher return will increase the probability that an individual will attend college. Likewise, higher ability is associated with increased likelihood of college attendance.

An Aggregate Model

A model of SAT participation and scores for an entire birth cohort can be developed from the above model of individual behavior. It is necessary to make some additional assumptions about the distribution of k_i^* in a cohort. For convenience, the distribution of is assumed to be normal. The distribution is given by:

$$k_i^* \sim N(\mu^*, \sigma_{11}) \tag{3.14}$$

where μ^* is the mean value of k_i^*:

$$\mu^* = \frac{1}{n}\sum_{i=1}^{n} k_i^* \tag{3.15}$$

and the individual departures from the cohort mean are defined as:

$$\epsilon_{1i} = k_i^* - \mu^* \tag{3.16}$$

It can readily be seen how the cohort mean will respond to changes in the college wage.

$$\frac{\partial \mu *}{\partial w_c} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial k_i^*}{\partial w_c} > 0 \tag{3.17}$$

Not every student who selects the college track takes the SAT, but it is assumed that all students who take the SAT choose the college track, and hence accumulate their optimal amount of college capital, k_i^*. Test-taking is thought to be related to several factors: the attractiveness of attending college, the degree to which the SAT is required of prospective students, and the type of school to which the individual will apply. As has been seen, the attractiveness of college attendance increases for a positive change in the relative college wage, while the degree to which the SAT is used by colleges would also have a positive influence on the probability that an individual will take the SAT. This might be reflected by the proportion of schools requiring the SAT.

The student's decision concerning SAT-taking can thus be modeled in the following fashion. A student will take the SAT if and only if:

$$z+\epsilon_{2i}\geq 0 \qquad (3.18)$$

The term z is defined as some increasing function of R and of the proportion of colleges using the SAT. These variables are observable, and are the same for every individual in a single cohort. The second term, ϵ_{2i}, is an individual disturbance which reflects differences among individuals, that is, ability. The relation between ability and ϵ_{2i} is thought to be positive, since students with higher abilities are expected to focus on more exclusive schools, and thus more likely to take the SAT. It is assumed that ϵ_{1i} and ϵ_{2i} are joint normally distributed.

$$\begin{pmatrix}\epsilon_{1i}\\ \epsilon_{2i}\end{pmatrix}\sim N\left(\begin{matrix}0\\0\end{matrix},\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & 1\end{bmatrix}\right) \qquad (3.19)$$

Individuals with higher ability are expected to have a higher value for ϵ_{2i} and thus to be more likely to take the SAT. Ability is also positively correlated with k_i^*. Thus, individuals with higher k_i^* are more likely to take the SAT, and σ_{12} is expected to be greater than zero.

The probability that a member of a cohort will take the SAT is given by

$$Prob(z+\epsilon_{2i}\geq 0)=\Phi(z) \tag{3.20}$$

where:

Φ = the standard normal distribution function.

Since the students who take the SAT are a self-selected sample, the expected level of college capital for an individual who does take the SAT will be conditional upon z and the individual disturbance. This can be described as follows:

$$\begin{aligned}\overline{k}_{SAT}&=E(k_i^*|z+\epsilon_{2i}\geq 0)\\&=\mu^*+E(\epsilon_{1i}|z+\epsilon_{2i}\geq 0)\\&=\mu^*+\sigma_{12}\ \lambda(z)\end{aligned} \tag{3.21}$$

where:

$$\lambda=\frac{\phi(z)}{\Phi(z)}>0 \tag{3.22}$$

and

ϕ = the standard normal probability density function.

A prediction of how the average observed SAT score is influenced by a change in the college wage, w_c, can now be made. The response of average SAT score to a change in the return to college is given by:

$$\frac{\partial \bar{k}_{SAT}}{\partial w_c} = \frac{\partial \mu^*}{\partial w_c} + \sigma_{12} \frac{d\lambda}{dz} \frac{\partial z}{\partial w_c} \tag{3.23}$$

It has already been seen that the first term is positive. Likewise, σ_{12} and the partial of z with respect to w_c are positive. Thus, the sign of the second term depends upon the sign of the partial of λ with respect to z. This is given as follows:

$$\frac{d\lambda}{dz} = -z\lambda - \lambda^2 < 0 \tag{3.24}$$

The second term of (3.23) is then negative. Although increasing returns to college education imply that those individuals taking the SAT will invest more heavily in those skills which are needed in college, the proportion of the cohort taking the exam will also increase. This selection effect is negative, and thus the sign of the total effect depends upon the relative magnitude of the two effects of an income change. Average SAT scores could either rise or fall in response to increasing returns to college, because the effect of rising returns in inducing more SAT-taking will have a negative effect on average scores, as more less-able students enter the test-taking population.

Discussion

Two implications of the above model may be important in explaining recent SAT trends. First, the relative college wage and the level of college capital acquired by students who choose the college track are positively correlated. As will be seen, the period of falling SAT scores corresponded closely with a downward trend in relative returns to college. Further, both trends reversed and began climbing at roughly the same time, with the upturn in the relative college wage preceding the upturn in SAT scores by one to three years. The model suggests that these score fluctuations might be explained in part by the changes in the return to higher education.

The second important implication of the model concerns the effects of self-selection of test-takers. As returns to college increase, so does the proportion of a cohort taking the exam. The effect on average scores of increased participation is clearly negative in this model. Other researchers have empirically examined this selection effect and concluded that it is negative. (For example, see Jackson, 1976; Turnbull,1985; Dynarski, 1987.) If this is so, the observed drop in SAT scores was less than it would have been in the absence of a selection effect, because cohort participation fell during the score decline. Similarly, increasing participation rates in recent years may be diminishing the measured recovery in scores. The relation between college returns and participation rates seems to follow the pattern predicted by the model.

The "explanation from economics", then, is comprised of three hypothesized effects. First, increasing economic returns to college are expected to lead to increased accumulation of intellectual capital among students attending college or taking the SAT. Second, increasing returns are thought to make college attendance and SAT-taking more attractive, inducing more members of a cohort to engage in these activities. Third, the effect on SAT scores of increased participation are hypothesized to be negative, as more less-able students are tested. How well do these hypotheses stand up to the empirical evidence?

CHAPTER 4

DATA AND ECONOMETRIC MODEL

Introduction

The data and econometric methods used in the study are discussed in this chapter. First, the data are described, and a number of variables which must be constructed are explained. The trends of certain key variables important to the hypotheses between 1967 and 1987 are briefly considered in order to give a picture of what occurred during the period. A basic econometric model is formulated, consisting of a system of equations representing SAT scores and proportions of cohorts taking the exam. Several possible specifications for the adjustment of scores over time are given, and estimation procedures are discussed.

The Data

Sources

The primary source for data on the SAT is the College Board, an organization of colleges which concerns itself with testing and college admissions issues. The College Board helped to found the Educational Testing Service and still maintains very close ties with it.

The College Board keeps a variety of statistics on the SAT and test-taking populations. Average scores and information on test populations and score distributions, broken down by test (math or verbal), sex, and year in school (junior or senior) are kept for each year. (The College Board has not, unfortunately, kept data on the makeup of the remainder of the SAT-taking population, i.e. freshmen, sophomores, and others.)

Also, in 1972 the ETS introduced the Student Descriptive Questionnaire (SDQ). This questionnaire, given to all SAT-takers as part of the SAT, is designed to give a general picture of the makeup of the SAT-taking population and asks each participant for (self-reported) information including grade point average, the type of coursework taken, extra-curricular activities, plans for future education and careers, and family size and income. Each year since 1972 the College Board has issued a report, "College Bound Seniors", based upon the data obtained from the SDQ. These reports include mean SAT scores, by test and sex, for all seniors in that year who participated at any time in the Admissions Testing Program (ATP), the College Board's name for its series of college entrance examinations.

In computing the scores for college bound seniors, the College Board includes only the most recent scores for those test-takers who have taken the test more than once. (Repeat test-taking seems to be a useful strategy in obtaining higher SAT scores.) For the years 1967 to 1971, the early years of the score decline, the College Board compiled estimates of the college bound senior scores. Because the average scores for these years were calculated by a different method than for later years, they may not measure exactly the same thing. Also, estimates of the numbers of college bound seniors were not calculated for the 1967 to 1971 period. The College Board also provided figures on institutional membership in the College Board.

Statistics on income and demographic trends, as well as additional information on trends in higher education were obtained from various publications of the U.S. government. An important data source is the Digest of Educational Statistics published by the U.S. Department of Education. Estimates of average yearly public and private tuitions in U.S. institutions of higher education, numbers of institutions of higher education, and numbers of high school graduates in each year were obtained from this document.

The U.S. Bureau of the Census series P-25, "Estimates of the Population of the United States by Age, Sex, and Race", was used to obtain figures on the numbers of seventeen-year-old males and females in the United States in each year covered by this study. The P-60 series, "Money Income of Households, Families and Persons in the United States", was the source for data on income. The present study made use of three income data series: median annual income for male high school graduates aged 25-34 working year round full-time, median annual income for males aged 25-34 with four years of college working year-round full-time, and median annual income for males aged 45-54 working year-round full-time.

Generating Variables

The principal hypothesis of this study is that educational choices are strongly influenced by economic returns to education and that this influence should be reflected in SAT participation and scores. The theoretical model of Chapter 3 indicates that the relative income of college graduates may play a key role in this regard. In order to measure this relative income, a variable, RELY, is constructed, by dividing the median income of American males aged 24-35 working year-round full-time (Y^{coll}) by the median income of American males with only a high school education working year-round full-time (Y^{hs}), and taking the natural logarithm of the resulting figure.

$$RELY_t = \ln\left(\frac{Y_{t-1}^{coll}}{Y_{t-1}^{hs}}\right) \quad (4.1)$$

The 25-34 age group is used because with discounting, the income earned in these years is weighted more heavily than income in later years. Also, to the extent that students might be myopic in their decision making, income prospects soon after leaving school will be especially important. Also, income for year-round full-time workers is used to avoid temporary effects -- such as those associated with the business cycle -- in the measure of relative income. Presumably, students are primarily concerned with relative permanent income.

It should be noted that for each year the income values used in generating RELY are those which pertain to the previous year. This is because test data are recorded by school year and dated by year of high school graduation, while income data is recorded by calendar year. Most graduating seniors who have taken the SAT did so in the spring or fall of the previous calendar year, that is, in the spring of their junior year or the fall of their senior year. Thus, the most recent information on incomes which students graduating in year t could be expected to have observed at the time they actually decided whether or not to take the SAT would be that for year t-1.

The SAT-taking population is self-selected, and therefore average test scores may be affected by factors which influence students to take the exam. An important implication of the theoretical model described in Chapter 3 is that when returns to education increase, the probability of a student deciding to take the SAT will also increase, thereby boosting cohort participation in the SAT program. According to the model, this increase in proportion of a cohort tends to be drawn from lower ability students. Thus the selection effect, i.e. the changes in the proportion of a cohort taking the exam, is expected to influence average scores.

Because the available data on the size of the test-taking population is in terms of raw numbers, it is necessary to scale the data in order to obtain the proportion of each cohort choosing to take the SAT. The numbers of male and female junior and senior test-takers are obtained from the College Board records for each year from 1967 to 1987. The same is done for male and female "college bound senior" (hereafter referred to as "cohort") test-takers for the years 1972-1987. These numbers are then scaled separately by the number of male and female graduating high school seniors in the respective cohorts.

For the senior and cohort data, the numbers for each year are simply divided by the number of high school graduates in that year. For the junior data, each observation is divided by the number of graduates in the following year, in order to scale the number of junior test-takers for each year by its own graduating cohort size. The form of the variables thus generated is as follows:

$$PROPMC_t=\frac{CM_t}{HSGM_t} \qquad PROPFC_t=\frac{CF_t}{HSGF_t}$$

$$PROPMS_t=\frac{SM_t}{HSGM_t} \qquad PROPFS_t=\frac{SF_t}{HSGF_t} \quad (4.2)$$

$$PROPMJ_t=\frac{JM_{t-1}}{HSGM_t} \qquad PROPFJ_t=\frac{JF_{t-1}}{HSGF_t}$$

where

CM_t = cohort male test-takers, year t;

CF_t = cohort female test-takers, year t;

SM_t = senior male test-takers, year t;

SF_t = senior female test-takers, year t;

JM_t = junior male test-takers, year t;

JF_t = junior female test-takers, year t;

$HSGM_t$ = male high school graduates, year t;

$HSGF_t$ = female high school graduates, year t;

Unfortunately, the unavailability of cohort (college bound senior) test-taker numbers for the years 1967 to 1971 makes it impossible to directly calculate the proportion of students taking the SAT in those years. These missing years are crucial for the estimation, as they include much of the early part of the score decline. Thus the first step is to obtain an estimate of these proportions for both males and females.

This is done by regressing the cohort proportions for the years in which they are available on the corresponding senior proportions, junior proportions, and year. The resulting coefficients are then used with the senior and junior data for 1967 to 1971 to generate predicted values for the proportion variables for those years. (A more complete description of the method employed and the results is found in Appendix B.)

The relative income derived from a college education would not be the only economic factor influencing college attendance. The immediate financial cost of attending college would also be relevant. To capture this cost the mean tuitions for U.S. public (TUITPUB) and private (TUITPRIV) 4-year institutions of higher education are averaged. Over the 1967 to 1987 time period the number of public and private institutions of higher education increased steadily, with the ratio remaining roughly 1:1.

In order to attempt to measure the importance of this nominal dollar figure, it is scaled by some measure of ability to pay. As a proxy for this, the median income of American males aged 45-54 working full-time year-round is used (MEDY), on the assumption that this gives an approximate measure of the abilities and resources of families and students to pay for higher education. The resulting variable, COST, can be expressed as follows:

$$COST_t = \frac{(TUITPUB_t + TUITPRIV_t)}{2MEDY_t} \tag{4.3}$$

Two caveats must be expressed concerning this variable. A possible weakness is that it may not necessarily reflect the actual financial outlay of students after student aid programs are considered. Also, it might be argued that tuition levels would be endogenous to any equation linking the proportion of a cohort attending college or taking the SAT with other factors. An increase in proportion might reflect an increase in the demand for college education, leading to higher tuition.

Another factor which would seem likely to play an important role in test-taking behavior is the degree to which the SAT is required of prospective entrants by colleges and universities. Unfortunately, data on the number of institutions requiring the SAT does not exist. As a proxy the number of institutions holding membership on the College Board (CB) was used, scaled by the number of institutions of higher education (#COLL), giving the variable MEMB.

$$MEMB_t = \frac{CB_t}{\#COLL_t} \tag{4.4}$$

A factor which is thought to have influenced college choices for males in the 1960's and early 1970's was the military conscription associated with the Viet Nam war, or, more precisely, the draft exemptions available to young men attending college. Such an inducement to enter college would be expected to have some impact on test-taking associated with college admission. The draft in these years selected from the male population between the ages of 18 and 26. However, most draftees were aged 18-20. A scaled measure of draft pressure is constructed by dividing the number of draftees in any given year, $DRAFT_t$, by the sum of the seventeen year old males (SEV_t) in that year and the two previous years. The resulting variable DDRAFT is as follows:

$$DDRAFT_t = \frac{DRAFT_t}{\sum_{i=t-2}^{t} SEV_i} \qquad (4.5)$$

Trends from 1967 to 1987

An overview of basic trends in the variables may help to give a feel for the nature of the SAT decline. It will be seen that trends in proposed explanatory variables seem to be closely correlated trends in SAT scores. The following initial examination of the data gives an idea of the basic trends occurring in the 1967-87 time period, and raises several interesting points. Also, a potential problem for estimation becomes apparent.

Cohort SAT scores exhibited a strong downward trend through the period. The high points for male math and female verbal scores came in 1967, while male verbal scores and female math scores peaked in 1968. By the early seventies all scores were plummeting, and in 1980 male cohort scores on both sections of the SAT reached their lowest points. Female math scores hit bottom in 1979 at 443 and remained there for four years. Female verbal scores reached their lowest point in 1981. The magnitude of the plunge ranged from 23 points on the male math test to 50 points on the female verbal test. For the entire decline (beginning in 1963) the total decline across sexes was .28 standard deviations on the mathematics exam and .48 standard deviations on the verbal.

Upon reaching their low points, all four scores hovered near their low points until the mid-eighties, when moderate improvements began. For both males and females the rise in math scores was noticeably stronger than the rise in verbal scores. (See Appendix A for a listing of scores and all other variables.)

The proportion of high school students taking the SAT was also changing during this period. For both males and females this figure reached a low point in 1977, after which it increased steadily.

These changes in the proportions of students taking the examination occurred at the same time that numbers of students graduating from high school reached an all-time high in the United States. Numbers of males graduating from high school peaked in 1976, while the number of females graduating and the total number of graduates peaked in 1977. All three numbers then began gradual declines to low points in 1986.

However, the decline in the proportion of students taking the SAT was not simply due to the growth of the student population. The actual number of male students taking the exam declined. Interestingly, the number of female test-takers was generally increasing through this period (except for a sharp drop in 1983 and 1984), while the number of males declined steadily until 1985.

The relative economic return to college education as measured by the RELY variable described above reached a peak in 1970 and then fell until 1978. During this period the earnings ratio for a worker with four years of college to that of a worker with only a high school education fell from 1.43 to 1.15 (for males, age 25-34, employed year-round full-time). This ratio began climbing slowly after 1978, but did not approach previous levels until 1985.

It should be noted that the trends in returns to education and the proportion of the cohort taking the exam appear to follow each other closely. This is in accord with the hypothesized effect of relative income on proportions of cohorts taking the SAT. Unfortunately, both are thought to be important explanatory variables for scores, and the possibility of multicollinearity arises. A high degree of correlation between RELY and the proportion variables could make it difficult to distinguish between the effects of college returns and selection in empirical estimation.

The number of members on the College Board in 1967 stood at 1017. Membership increased to 1190 in 1973, after which it fluctuated until the early eighties. After reaching a low of 1126 in 1983, it then rose until in 1987 there were 1166 members. During this time the number of institutions of higher learning in the United States rose steadily, from 2532 in 1967 to 3406 in 1987. Thus, the trend in the scaled College Board variable, MEMB, was downward. If this roughly reflects the unobserved trend in the degree to which the SAT was required of students, the effect on the proportion of the cohort taking the test would be expected to be negative, as test-taking is thought to be positively correlated with the use of examination results by colleges.

Econometric Models

The principal hypothesis is that the relative returns to college education should affect both the probability of students choosing to take the SAT and the skill levels (scores) of the individual test-takers. The potentially observable effects would be seen, if they exist, in average SAT scores and in proportions of cohort participation. Thus, the general model consists of a system of six equations; one for the male cohort proportion, one for female cohort proportion, and one for each of the SAT scores: male mathematics, male verbal, female mathematics, and female verbal.

The score equations would not be expected to adjust fully in the time period in which an income change occurs. Student preparation for the SAT is thought to take place over a period of time; acquiring the math, verbal, and reasoning skills thought necessary to do well on the SAT would take place throughout the high school years, at the very least. Thus, the effects of factors on student study habits would not be expected to be fully reflected in test scores until after a period of several years.

For the equations attempting to capture the factors influencing the proportion choosing to take the exam, there is no a priori reason to expect an incomplete contemporaneous adjustment. The decision to take or not take the SAT takes place at one moment in time, relatively shortly before the exam, and might be expected to respond immediately to the latest available information regarding incentives to take the exam and/or attend college.

Explanatory variables for the proportion equations would be those factors likely to be important in college-choice decisions, and would include returns to college education, costs of such education, and -- for males before 1972 -- the risk of military conscription. The decision to take the SAT would also be expected to be influenced by the degree to which the examination is required by colleges. In order to capture this effect, the MEMB variable described above is included.

Numerous researchers have argued for the existence of family size and other demographic effects on the intellectual development of children, and as previously discussed several studies have found empirical confirmation of such effects. One result of the theoretical model of Chapter 3 is that a student's level of "ability" might well influence his decision to pursue further education, or to take the SAT. In an attempt to catch such demographic effects, average birth order (ABO) is used as an exogenous explanatory variable. Because each birth cohort enters the senior year of high school 17 years after birth, the appropriate ABO figure for the college bound senior cohort in year t is the U.S. average birth order in year t-17. It should be cautioned that this variable may well pick up a variety of interrelated family and cohort size effects, and might be expected to overestimate any effect due simply to birth order. Also, as discussed above, previous studies have found that birth order effects are frequently overestimated in time series studies.

It is also clear that changes related to time have been occurring in education and career choices, especially for females. The increasing presence of women in the workplace, and especially in male-dominated occupations, would likely affect SAT-taking behavior. In order to pick up such trends, a time trend variable, YEAR, is included in each proportion equation.

In accordance with these theoretical concerns, a general form for the equations relating the proportion of a cohort taking the SAT to possible explanatory factors is suggested. These proportion equations are of the form

$$\begin{aligned} PROPMC &= f(RELY, COST, MEMB, ABO, YEAR, DDRAFT) \\ PROPFC &= g(RELY, COST, MEMB, ABO, YEAR) \end{aligned} \qquad (4.6)$$

where f and g are arbitrary functions. The difference between the two equations is the presence of the draft variable in the equation for males.

It has been argued that the score equations would not be expected to show full contemporaneous adjustment to changes in explanatory variables. The general form of the equations used in this study is

$$SCORE = h(RELY, ABO, YEAR, PROPMC \text{ or } PROPFC) \qquad (4.7)$$

where the inclusion of lagged dependent variables or lagged relative return variables to account for the dynamics of adjustment of preparation patterns to changing incentives is implicit.

The presence of RELY in the score equations would be expected to pick up the hypothesized effect of economic incentives on student preparation for the SAT. The COST variable is not included here, for two reasons. First, the model presented in Chapter 3 implies that once the decision has been made to attend college the direct financial costs do not play any further role. Also, initial preliminary regressions failed to find any significant role for COST in the score equations. With neither a priori nor empirical reason to include COST, it is excluded here.

An important question in inquiries into the causes of the SAT score fluctuations is the degree to which changing proportions of the student population taking the exam affect the outcomes. A prediction of the theoretical model is that given non-homogeneous students and some distribution of intellectual talent, it seems likely that as the proportion of a cohort taking the exam increases, the marginal student is increasingly less able, which in turn would tend to lower the average score, ceteris paribus. It has already been explained why it is of crucial importance to recognize the existence of such selection effects in studies of phenomena such as the SAT where the study population is self-selected. In order to account for such self-selection effects, a variable based on the proportion of students choosing to take the exam must be included. The variables used are described below.

Demographic effects, especially those relating to family size, are expected to be important in affecting measurements of intellectual achievement such as the SAT. In order to account for this, the ABO variable described above is included in the score equations. Again, the caveat concerning interrelated demographic effects and possible overestimation applies.

A time trend variable, YEAR, is also included in the score equations. As with the proportion equations, changes in education and career patterns through time unrelated to the above factors may influence intellectual achievement and test scores.

For both the proportion and score equations, the dependent variables in the first five observations (1967-71) are estimates. In order to account for any consistent bias in the estimates, a dummy variable (D71) is included in each of the equations. The value of the dummy is set equal to one for the years 1967 through 1971, and zero for the remaining years.

The dependent variables for the proportion equations are derived directly from the PROPMC and PROPFC variables described above. The proportion of a cohort taking the SAT will be some fraction between one and zero. This number is transformed to the corresponding critical value in the standard normal distribution, a number which is not bounded. This is done for both males and females, giving the variables CRITM and CRITF.

$$CRITM_t = F^{-1}(PROPMC_t)$$
$$CRITF_t = F^{-1}(PROPFC_t) \qquad (4.8)$$

For the score equations, the cohort proportion variable designed to capture the selectivity effect is as follows. For each observation, the CRITM or CRITF variable is used as the argument in a ratio consisting of the standard normal probability density function over the cumulative density function, the λ function of the theoretical model. The resulting variables, LAMM for males and LAMF for females, should capture selectivity effects in the SAT score equations[1].

$$LAMM_t = \frac{f(CRITM_t)}{F(CRITM_t)}$$

$$LAMF_t = \frac{f(CRITF_t)}{F(CRITF_t)} \qquad (4.9)$$

The model will be estimated as a system of six equations; female cohort proportion, male cohort proportion, female math, female verbal, male math, and male verbal. Because LAMM and LAMF in the score equations are transformations of the dependent variables in the proportion equations, they are endogenous explanatory variables. If the disturbances in the score and proportion equations are correlated, the LAM variables will be correlated with the error terms in the score equations, in which case OLS estimates are biased and inconsistent.

To overcome this difficulty, instruments are constructed. These are variables which are highly correlated with LAMM and LAMF, but not with the score

---

[1]These particular functional forms for the proportion and selectivity variables are those implied by the assumption that optimal capital accumulation, k_i^*, and propensity to take the SAT, $z + \epsilon_{2i}$, are joint normally distributed. If this distribution is not in fact joint normal, then some other functional forms would be appropriate. However, the variables would still have similar properties; CRIT and LAM would remain monotonically increasing and decreasing functions of the proportion taking the SAT.

disturbances. The instruments, LAMMHAT and LAMFHAT, are formed by first obtaining predicted values for CRITM and CRITF. These values are derived by estimating the proportion equations as a system using the Seemingly Unrelated Regressions method. The predicted values for CRITM and CRITF resulting from this procedure are then transformed by the same method described above, that of dividing the appropriate PDF by the corresponding CDF. The instruments so obtained are asymptotically uncorrelated with the score equation disturbances.

As discussed above, there also exists the potential problem of multicollinearity between the proportion variables, LAMM and LAMF, and the relative income variable, RELY, and perhaps other right hand side variables. Multicollinearity problems can lead to imprecise estimates, difficulties in rejecting hypotheses, and sign problems. One method for dealing with multicollinearity problems is to utilize outside information to derive estimates for the coefficients for one of the linearly related variables. Such outside information exists for the proportion variables. As will be seen, values for the LAMM and LAMF coefficients derived from the literature may prove useful.

The outside information is derived from the previously discussed work of Dynarski (1987). Using a pooled cross-sectional time series approach with SAT data for separate states, Dynarski estimated that at a .27 mean value for SAT participation in the U.S. (the value for 1981), a one percent increase in participation corresponds to a 1.8 point decrease in average SAT score. Dynarski obtained separate estimates for the effects of participation rates on both the verbal and mathematical sections of the SAT. His results are as follows:

$$SAT_v = -16.10p^* + X\underline{\beta}_v$$

$$SAT_m = -20.02p^* + X\underline{\beta}_m \; ; \qquad (4.10)$$

$$p^* = \ln\left(\frac{p}{1-p}\right)$$

where

p = participation rate,

X = a matrix of school input and socioeconomic variables thought to be related to student achievement,

$\underline{\beta}$ = a vector of estimated coefficients.

From Dynarski's results the following relationships may be obtained:

$$\frac{\partial SAT_v}{\partial p} = -16.10\frac{\partial p^*}{\partial p}$$

$$\frac{\partial SAT_m}{\partial p} = -20.02\frac{\partial p^*}{\partial p} \qquad (4.11)$$

$$\frac{\partial p^*}{\partial p} = \frac{1}{p(1-p)}$$

Evaluated at the mean of p = .27, the effects of an increase in SAT participation are thus given by:

$$\frac{\partial SAT_v}{\partial p} = -81.684$$

$$\frac{\partial SAT_m}{\partial p} = -101.57 \qquad (4.12)$$

These results may be used in the present model to derive estimates of the coefficients for LAMM and LAMF in the score equations. The general form for a score equation can be written as:

$$SAT = \alpha\lambda + Z\underline{\delta} \qquad (4.13)$$

where

λ = the LAMM or LAMF variable,

Z = a matrix of the remaining explanatory variables,

α = the coefficient of λ, and

$\underline{\delta}$ = a vector of coefficients on Z.

The proportion variable λ can also be expressed as

$$\lambda = \frac{f(CRIT)}{F(CRIT)} = \frac{f[F^{-1}(p)]}{p} = \frac{f(y)}{p}\ ; \qquad (4.14)$$

$$y = F^{-1}(p)$$

From this it is clear that the response of SAT scores is given by:

$$\frac{\partial SAT}{\partial p} = \alpha\frac{\partial \lambda}{\partial p} \qquad (4.15)$$

where

$$\frac{\partial\lambda}{\partial p} = -\frac{f(y)}{p^2} + \frac{1}{p}\cdot\frac{df(y)}{dy}\cdot\frac{dy}{dp}$$

$$\frac{dy}{dp} = \frac{1}{f(y)} \tag{4.16}$$

$$\Rightarrow \frac{d\lambda}{dp} = -\frac{f(y)}{p^2} - \frac{y}{p}$$

When the rate of change of λ with respect to proportion is evaluated at p = .27, the result is thus:

$$\frac{d\lambda}{dp} = -2.266. \tag{4.17}$$

Dynarski found that participation effects were different for the two sections of the SAT. The results are as follows:

$$\frac{\partial SAT_v}{\partial p} = -81.684 = \alpha_v(-2.266);\quad \alpha_v = 36.04$$

$$\frac{\partial SAT_m}{\partial p} = -101.57 = \alpha_m(-2.266);\quad \alpha_m = 44.82 \tag{4.18}$$

These are the Dynarski values for the coefficients of LAMM and LAMF. It will be seen that these coefficients may be useful in improving estimates of the model. Using outside information for the LAM coefficients will circumvent possible collinearity problems.

The model is estimated as a system of six equations, as described above. Two different dynamic specifications were chosen for the score equations. The first includes two lagged values of the relative income variable, RELY, in addition to the current value. Thus the impact of a change in relative returns to college in year t affects scores not only in year t but also in years t+1 and t+2. In this way the model may capture the process of slow adjustment in acquisition of intellectual capital by students. The full effect of the change in RELY is not manifested until the third time period following the change. The full effect of a one-time, permanent change in RELY is the sum of the coefficients times the amount of the change. It is expected that all of the estimated coefficients for the RELY variables would be positive, indicating a positive effect of economic incentives on student accumulation of intellectual capital. This model, which will be referred to as the 3RELY model, has the following specification for the system.

$$MMC_t = C_1 + \alpha_1\ RELY_t + \beta_1\ RELY_{t-1} + \gamma_1\ RELY_{t-2} + \delta_1\ ABO_t + \epsilon_1\ YEAR_t + \zeta_1\ LAMM_t + {}^{n}{}_{1}\ D71_t + u_{1t}$$

$$VMC_t = C_2 + \alpha_2\ RELY_t + \beta_2\ RELY_{t-1} + \gamma_2\ RELY_{t-2} + \delta_2\ ABO_t + \epsilon_2\ YEAR_t + \zeta_2\ LAMM_t + \nu_2\ D71_t + u_{2t}$$

$$MFC_t = C_3 + \alpha_3\ RELY_t + \beta_3\ RELY_{t-1} + \gamma_3\ RELY_{t-2} + \delta_3\ ABO_t + \epsilon_3\ YEAR_t + \zeta_3\ LAMF_t + \nu_3\ D71_t + u_{3t}$$

$$VFC_t = C_4 + \alpha_4\ RELY_t + \beta_4\ RELY_{t-1} + \gamma_4\ RELY_{t-2} + \delta_4\ ABO_t + \epsilon_4\ YEAR_t + \zeta_4\ LAMF_t + \nu_4\ D71_t + u_{4t}$$

$$CRITM_t = C_5 + \alpha_5\ RELY_t + \theta_5\ MEMB_t + \iota_5\ COST_t + \delta_5\ ABO_t + \epsilon_5\ YEAR_t + \kappa\ DDRAFT_t + \nu_5\ D71_t + u_{5t}$$

$$CRITF_t = C_6 + \alpha_6\ RELY_t + \theta_6\ MEMB_t + \iota_6\ COST_t + \delta_6\ ABO_t + \epsilon_6\ YEAR_t + \nu_6\ D71_t + u_{6t}$$

where

MMC_t = average cohort SAT score, math, males,

VMC_t = average cohort SAT score, verbal, males,

MFC_t = average cohort SAT score, math, females,

VFC_t = average SAT score, verbal, females,

C_i = constant term, i^{th} equation,

u_{it} = error term, i^{th} equation,

i = (1,...,6), and

the subscript t refers to the year of the observation,

t = (1967,...,1987).

A second form for the score equations includes a lagged dependent variable instead of lagged values of RELY. In this specification, the LDV model, the scores in year t are functions not only of relative income and other variables in some year t, but also of scores in the previous year. In this way the effect of a change in RELY (or any other explanatory variable) in year t is not fully manifested in year t, but continues to have effect through time as scores adjust. The estimated coefficients for the lagged dependent variables then are expected to be positive and less than one. (A specification using a lagged dependent variable and lagged RELYs was also tried, but the estimated coefficient on the lagged dependent variable proved insignificantly different from zero.) The proportion equations in the LDV model are identical to those of the 3RELY model. The LDV score equations are:

$$MMC_t = C_1 + \xi_1 MMC_{t-1} + \alpha_1 RELY_1 + \delta_1 ABO_t + \epsilon_1 YEAR_t + \zeta_1 LAMM_t + \nu_1 D71_t + u_{1t}$$

$$VMC_t = C_2 + \xi_2 VMC_{t-1} + \alpha_2 RELY_2 + \delta_2 ABO_t + \epsilon_2 YEAR_t + \zeta_2 LAMM_t + \nu_2 D71_t + u_{2t}$$

$$MFC_t = C_3 + \xi_3 MFC_{t-1} + \alpha_3 RELY_3 + \delta_3 ABO_t + \epsilon_3 YEAR_t + \zeta_3 LAMF_t + \nu_3 D71_t + u_{3t}$$

$$VFC_t = C_4 + \xi_4 VFC_{t-1} + \alpha_4 RELY_4 + \delta_4 ABO_t + \epsilon_4 YEAR_t + \zeta_4 LAMF_t + \nu_4 D71_t + u_{4t}$$

(Note that the cohort scores do not exist for years before 1967. This reduces the number of available observations for the LDV model by one.)

According to the hypothesized influence of economic incentives on SAT scores, the estimated coefficients for the RELY variable in the proportion equations (equations 5 and 6) should be positive. Also, the hypothesized effect of RELY on the proportion equations should be further manifested in the score equations through LAMM and LAMF, the proportion variables. As the proportions taking the exam increase, the scores are expected to decrease. This would translate into positive values for the LAMM and LAMF coefficients.

The estimated values for the coefficients of MEMB in the proportion equations are also expected to be positive. As discussed above, MEMB serves as a proxy for the degree to which the SAT is required of prospective college students.

Changes in MEMB and changes in the proportion taking the exam should be positively related.

In all equations the coefficients for the ABO variable are expected to be negative. Family size effects on intellectual achievement have been documented as being negative with increasing family size.

The variable capturing the effects of the draft and college deferments, DDRAFT, in the male proportion equation is expected to have a positive coefficient. As the proportion of males being conscripted from the draft age population increases, the perceived probability of being drafted should increase. This increased perceived risk should in turn increase the demand of males for college (and deferments) and thus for the SAT.

The time trend variable, YEAR, has no particular sign or expected value dictated by the working hypotheses. However, given that the general trend of SAT scores through the 1967-87 period was downward, and that there are doubtlessly a variety of unknown factors which influenced the downturn in scores, YEAR might be expected to capture the effects of these unknown factors and therefore be negative. Likewise, there is no a priori expectation for YEAR's effect in the proportion equations. In the early portion of the study period the general trend was downward, while in later years the trend was upward. For females the decline appears to have been less pronounced, and the increase sharper. From this it would seem likely that the value of the coefficient for YEAR in the female proportion equation would be positive and greater than the coefficient in the males' equation.

CHAPTER 5

EMPIRICAL RESULTS

Introduction

The results of the econometric estimation procedures are reported for the 3RELY and LDV models. The first estimations are of the models with no restrictions imposed on any coefficients. Next, the estimation is of the models with the coefficients on some of the economic variables set equal across exams and sexes in the score equations, and across sexes in the proportion equations. Finally, the results of estimation of the restricted models with the added restriction of the Dynarski values imposed on the LAMM and LAMF coefficients are reported, and another specification suggested by the results is estimated and examined.

The procedure which will be used for discussing the results is as follows: the results for the two models, 3RELY and LDV, will be considered together. The estimated coefficients will be treated in this order: first RELY in the proportion equations, then RELY in the score equations, lag structure, LAMM and LAMF, and finally the remaining variables and other results. All reported statistical tests are conducted at the .05 level of significance unless otherwise specified.

Unrestricted Models

The two models are initially estimated without imposing any restrictions. As discussed in Chapter 4, these are simultaneous equation systems, because the LAMM and LAMF variables in the score equations are transformations of the dependent variables in the proportion equations, and thus are endogenous explanatory variables. In addition, the disturbances are likely to be correlated across equations. For these reasons, the three stage least squares procedure (3SLS) is used, substituting as instruments for LAMM and LAMF the variables described above, LAMMHAT and LAMFHAT. The unrestricted 3SLS regressions for the 3RELY and LDV models give the results reported below in Tables 1 and 2.

For the proportion equations, the estimated coefficients of RELY are all positive, as predicted by theory. However, the hypothesis that the true coefficient is zero cannot be rejected for any of the estimates at the .05 level.

In the score equations RELY also seems to have a positive effect, but again the estimates are imprecise. For the LDV model, RELY has positive and statistically significant coefficients for the male and female verbal

Table 1. Results for 3RELY model, unrestricted.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2460.5 (6.13)	3493.2 (6.29)	3245.6 (4.81)	6186.8 (9.81)	-7.1671 (-.918)	-24.164 (-3.9)
$RELY_t$	7.31 (.49)	32.52 (1.58)	-1.67 (-.09)	28.85 (1.61)	.27442 (1.68)	.24498 (1.77)
$RELY_{t-1}$	-4.59 (-.32)	-21.91 (-1.09)	-13.46 (-.80)	-8.30 (-.48)		
$RELY_{t-2}$	-2.18 (-.16)	31.83 (1.72)	54.53 (3.30)	16.06 (1.00)		
YEAR	-.95 (-4.73)	-1.51 (-5.44)	-1.36 (-4.23)	-2.85 (-9.49)	.00376 (.99)	.01226 (4.05)
ABO	-15.75 (-2.05)	-9.63 (-.90)	-1.93 (-.23)	-12.93 (-1.67)	-.29799 (-3.86)	-.22335 (-3.53)
LAMM	-42.23	-49.36	-101.91	-78.56		
LAMF	(-1.54)	(-1.29)	(-2.51)	(-2.08)		
D71	-4.15 (-1.27)	-4.34 (-.95)	-3.99 (-1.02)	-2.42 (-.67)	.11890 (4.47)	.10321 (4.90)
MEMB					.89132 (1.20)	.75087 (1.26)
COST					-2.0718 (-2.34)	-1.8308 (-2.49)
DDRAFT					1.2808 (4.47)	
R^2	.9679	.9789	.9643	.9922	.9778	.9615
Durbin-Watson	2.35	2.04	2.39	2.74	2.64	2.53
degrees freedom	13	13	13	13	13	14

Table 2. Results for LDV model, unrestricted.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2502.2 (4.06)	5741.0 (5.96)	2897.2 (2.38)	6351.0 (4.42)	-4.3331 (-.46)	-21.113 (-2.85)
$Score_{t-1}$	-.04 (-.20)	-.37 (-2.06)	.12 (.61)	-.02 (-.13)		
RELY	8.61 (.65)	40.64 (2.01)	18.30 (.94)	36.72 (2.18)	.20126 (1.22)	.17294 (1.27)
ABO	-16.58 (-1.97)	-19.94 (-1.58)	-32.53 (-3.09)	-19.08 (-2.12)	-.33542 (-3.91)	-.26175 (-3.82)
LAMM LAMF	-29.62 (-.92)	-66.32 (-1.53)	26.75 (.50)	-51.13 (-1.04)		
D71	-2.60 (-.73)	-9.57 (-1.89)	.93 (.19)	-1.92 (-.47)	.11805 (.46)	.10347 (5.11)
MEMB					.75880 (.88)	.59872 (.88)
COST					-1.9881 (-2.25)	-1.7604 (-2.45)
DDRAFT					1.3354 (3.23)	
YEAR	-.96 (-3.72)	-2.54 (-5.71)	-1.24 (-2.24)	-2.94 (-4.36)	.00241 (.53)	.01080 (3.01)
R^2	.9627	.9686	.9646	.9910	.9778	.9668
degrees freedom	13	13	13	13	12	13

exams. These coefficients are also of sufficient magnitude to represent important influences on test scores. The RELY coefficients in the math equations are also positive, but much smaller, and for both the hypothesis that the coefficient is zero can't be rejected.

In the 3RELY model, the sums of the RELY coefficients in each equation are positive. However, six of the Table estimated coefficients are negative, and only one coefficient -- the large positive estimate for $RELY_{t-2}$ in the female math equation -- is significantly different from zero. The presence of the negative coefficients is unexpected. All of the coefficients for $RELY_{t-1}$ are less than zero. If the estimates are accurate reflections of reality, this implies an odd pattern of student response to changes in relative college returns. An increase in RELY in year t induces a positive response in scores for students who are seniors in high school, and a negative one for juniors. Of course, these negative estimates are not significantly different from zero. And with the exception of the male math equation, the sums of the RELY coefficients are sufficiently large that RELY would play an important positive role in determining scores.

Tests of the hypothesis that the coefficients for the lagged values of RELY, considered together, are equal to zero give mixed results. In separate tests for each score equation, this hypothesis can be rejected at the five percent level only for the female math equation. In a joint test of the lagged RELYs from all four equations the hypothesis cannot be rejected. (See Table 3 for a complete listing of tests and results from the unrestricted estimations.)

Table 3. Tests conducted with unrestricted model estimates.

Hypothesis	Statistic	Result, .05 level
3RELY model		
Coefficients for lagged RELYs each are zero, male math equation	F(2,79) = .12	Can't reject
Coefficients for lagged RELYs each are zero, male verbal equation	F(2,79) = 1.52	Can't reject
Coefficients for lagged RELYS each are zero, female math equation	F(2,79) = 5.72	Reject
Coefficients for lagged RELYS each are zero, female verbal equation	F(2,79) = .50	Can't Reject
All coefficients for lagged RELYs are zero	F(8,79) = 1.84	Can't reject
Sum of RELY coefficients is zero, male math	t(79) = .04	Can't reject
Sum of RELY coefficients is zero, male verbal	t(79) = 2.06	Reject
Sum of RELY coefficients is zero, female math	t(79) = 1.87	Reject
Sum of RELY coefficients is zero, female verbal	t(79) = 1.87	Reject

Table 3, continued.

Hypothesis	Statistic	Result, .05 level
3RELY model		
Sum of RELY coefficients is zero, each equation	$F(4,79) = 2.17$	Can't reject
Coefficients for $RELY_t$, $RELY_{t-1}$, $RELY_{t-2}$, and D71 are same across score equations, and coefficients of RELY, COST, and D71 same across proportion equations	$F(11,79) = 1.67$	Can't reject
Same across equations as above, with LAMM and LAMF coefficients equal to values from Dynarski	$F(19,79) = 3.50$	Reject
LDV model		
Coefficients for $Score_{t-1}$, RELY and D71 are same across score equations, and coefficients for RELY, COST, and D71 are same across proportion equations	$F(12,77) = 1.60$	Can't reject
Same across equations as above, with LAMM and LAMF coefficients equal to values from Dynarski	$F(16,77) = 2.31$	Reject

However, testing the hypothesis that RELY has no permanent or long run effect on the scores (i.e., that the sum of the coefficients on the three RELY variables in each equation is zero) results in rejection for one of the equations at the five percent level and two others at the ten percent level. Even so, a joint test of this hypothesis for all four equations fails to reject, apparently due to the effect of the male math equation estimates.

The coefficients of the lagged dependent variables in the LDV score equations are also unexpected. Three of the four have negative signs, and one of these is a statistically significant estimate. For the only positive coefficient (female math), the hypothesis of difference from zero can't be rejected. Again, these negative adjustment coefficients imply an unexpected pattern of adjustment.

Also surprising are the coefficients for the LAMM and LAMF variables. With one exception, the coefficients are negative numbers and of sufficient magnitude to imply a substantial effect for participation rates on scores. The LAM transformation results in variables which are themselves negatively related to participation; these results imply that increased participation results in higher scores. Two of the negative estimates are statistically significant.

Looking again at the proportion equations it can be seen that the estimated coefficients for COST are significant in all four equations, and the magnitude and signs indicate a strong negative effect, quite in accord with theory. Likewise, the effect of military conscription and college deferments as represented by the coefficients on DDRAFT in the male proportion equations is very strong for both models, and the estimates are significant.

The estimates of the ABO coefficients in all of the proportion and score equations are negative, as predicted. The estimates for these coefficients are significant in all of the proportion equations and in four of the eight score equations, appearing to corroborate previously documented negative effects on intellectual development for birth order and family size.

The coefficient for YEAR is significant for all equations in both models, except for the male proportion equations. The values are negative in all score equations, apparently indicating the presence of some important omitted variable(s) operating through time.

In both models the estimated values of the D71 coefficients were negative in the score equations, with one exception, and positive in the proportion equations. Significant values were found in three proportion equations and one of the score equations (a negative value). This suggests that the values calculated in this study (see Chapter 4) for the proportions for 1967 to 1971 are higher than what would be observed had they been calculated using the same procedures as for 1972 to 1987.

The hypothesis of an important role for economic returns to education in SAT-taking is not rejected in these estimates, but the hypothesized negative effect of increased participation seems extremely doubtful. However, the possible multicollinearity problems between RELY and the LAMM and LAMF variables might be responsible for these results. As discussed previously, the bringing of "outside" information into the models could overcome this problem. Also, the small number of observations do not allow for as precise estimation as might be liked. It is known that imposing a true restriction upon a model will improve the efficiency of an estimate. It does not seem unreasonable to speculate that SAT-taking behaviors are the same across exam, across sex, or across both. If this is so, imposing such restrictions would be of definite value in estimating the models.

With these points in mind a number of tests conducted with the estimated coefficients are reported. First, tests of the values for the LAM variables derived from Dynarski are tried for both models. In both cases the values are strongly rejected. Multicollinearity problems should result in large standard errors for the estimated values, making rejection of alternative hypotheses (such as that the estimated values are not significantly different from the Dynarski values) difficult. The rejection of the Dynarski values in both the 3RELY and LDV models is strong evidence against them.

Tests are also conducted with both models of a set of restrictions across sex and exam, in which the coefficients of the economic variables, the dummy variables, and the lagged dependent variable are restricted to common values. These restrictions correspond to the hypothesis that males and females respond in the same way to changes in these factors, and that students' preparation also responds the same in terms of both mathematical and verbal skill acquisition. The exact specifications and results of the tests are given in Table 3.

Testing of these restrictions in the 3RELY model estimates leads to a failure to reject -- it cannot be concluded that the restrictions are false. However, when these restrictions are tested in conjunction with the Dynarski values for LAMM and LAMF, they are rejected.

The results of the tests with the LDV model are quite similar. The restrictions across sex and exam cannot be ruled out, while testing of the hypothesis that both the restrictions and the Dynarski values are true does lead to rejection. The failure to reject the first set of restrictions in both models suggests the strategy of estimating the restricted forms of the models.

Restricted Models

The method of estimation for the restricted models is identical to that of the unrestricted models, with the only difference being the imposition of equality on certain coefficients across equations. For the proportion equations, the coefficients for RELY, COST, and D71 are restricted to be the same for males and females. For the score equations, the coefficients for RELY and D71 are set equal for every exam, as are the coefficients for the lagged values of RELY and the lagged dependent variable for the 3RELY and LDV models respectively. These restricted models are estimated, again using the 3SLS procedure. The results are reported in Tables 4 and 5.

For the proportion equations, RELY again has a positive coefficient of sufficient magnitude to make it important as a determinant of participation rates. However, the estimates are significantly different from zero only for the 3RELY model.

RELY also has a positive influence in the score equations. In the 3RELY model, the coefficients for $RELY_t$ and $RELY_{t-2}$ are positive, and the sum of all the coefficients is of sufficient size to imply an important role in score levels. Again, the unexpected negative sign on the coefficient for $RELY_{t-1}$ raises questions about the pattern of students' responses to changing incentives. Only the positive coefficient for $RELY_{t-2}$ is significantly different from zero. And although the sum of the RELY coefficients is positive, it is not significantly different from zero. (Tests are reported in Table 6.)

For the LDV model, the RELY coefficient is smaller, and the estimate is not significant. This does not indicate an important role for economic incentives.

Table 4. Results for 3RELY model, restricted.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	1868.9 (4.94)	3509.4 (8.40)	3873.5 (9.52)	6738.5 (16.77)	-8.630 (-1.22)	-24.169 (-4.30)
$RELY_t$	7.39 (.66)				.22837 (2.00)	
$RELY_{t-1}$	-10.39 (-.96)					
$RELY_{t-2}$	20.09 (1.97)					
YEAR	-.65 (-3.45)	-1.50 (-7.17)	-1.67 (-8.52)	-3.10 (-16.1)	.00437 (1.26)	.01219 (4.43)
ABO	-2.89 (-.40)	-12.37 (-1.54)	-11.22 (-1.51)	-9.18 (-1.31)	-.28563 (-5.02)	-.21513 (-4.03)
LAMM LAMF	-69.58 (-3.17)	-70.76 (-3.01)	-94.18 (-3.08)	-137.13 (-4.87)		
D71	-5.67 (-2.36)				.10097 (5.48)	
MEMB					1.2599 (1.95)	.91585 (1.71)
COST					-1.3796 (-2.22)	
DDRAFT					1.8217 (7.01)	
R^2	.9484	.9736	.9594	.9877	.9771	.9606
Durbin Watson	2.24	1.89	2.07	2.59	2.26	2.30
degrees freedom	28	28	28	28	28	29

Table 5. Results for LDV model, restricted.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2316.8 (5.68)	4177.9 (6.17)	3891.1 (6.19)	6581.9 (7.09)	-3.477 (-.43)	-18.788 (-2.76)
$Score_{t-1}$	-.0006 (-.57)					
RELY	13.73 (1.33)				.15474 (1.29)	
ABO	-12.43 (-1.89)	-25.36 (-2.81)	-29.06 (-3.47)	-20.39 (-2.70)	-.35202 (-5.69)	-.28284 (-4.87)
LAMM LAMF	-34.04 (-1.70)	-29.12 (-1.25)	-23.34 (-.80)	-83.58 (-2.92)		
D71	-3.26 (-1.42)				-.09756 (5.44)	
MEMB					.83623 (1.07)	.44762 (.71)
COST					-1.7843 (-2.91)	
DDRAFT					1.8596 (6.84)	
YEAR	-.88 (-4.87)	-1.84 (-5.82)	-1.69 (-5.91)	-3.04 (-6.99)	.00211 (.49)	.00969 (2.9)
R^2	.9618	.9694	.9628	.9892	.9970	.9667
degrees freedom	25	25	25	25	24	25

The lagged dependent variable coefficient is very small, and thus indicates little adjustment effect. Furthermore, it is negative, and insignificantly different from zero. This is not in accord with the hypothesis that student achievement responds over time to changes in

Table 6. Tests conducted with estimates from restricted 3RELY model.

Hypothesis	Statistic	Result, .05 level
Coefficients for lagged values of RELY are each zero	F(2,94) = 1.94	Can't reject
Sum of the RELY coefficients is zero	t(94) = 1.50	Can't reject

variables such as relative income. And it seems to be contradicted by the findings of the 3RELY model, where student response to changes in RELY occurs well after the change.

The estimated coefficients for LAMM and LAMF are all negative, and equality to zero can be rejected for five of eight of them. This is strong evidence against the proposed selectivity effect.

YEAR and ABO are significant and negative for the score equations in both models. Additionally, ABO plays an important negative role in the proportion equations for both models.

The estimated coefficients for the MEMB variables are positive and significant in the 3RELY model, and positive but insignificant in the LDV model. COST is strongly significant for both models, and negative, as predicted. DDRAFT was strongly positive and significant for males, while YEAR in the proportion equations is significant only for females, and positive.

The economic hypotheses of a positive influence for RELY on proportion, and a positive direct effect on score, appear to be defensible, given these results. The proposed selectivity effect, on the other hand, seems to be completely refuted by the empirical results. Oddly enough, it is this last hypothesis which has been previously documented. As there does exist strong reason to believe that the effect of an increase in the proportion of students taking the SAT would be negative, it may not be unreasonable to impose the Dynarski values on the restricted forms of the 3RELY and LDV models. Still, it should not be forgotten that these values are rejected in tests conducted with the data.

Restricted Models with Dynarski Values

The models estimated here are identical to the fully restricted models above, with the additional restriction that the coefficients on the LAMM and LAMF variables are set equal to the values calculated from Dynarski's findings. With values of these coefficients dictated by outside information rather than being estimated within the system, the problem of endogeneity in the score equations does not arise. The appropriate estimation procedure then is that of Seemingly Unrelated Regressions (SUR). The estimations give the results listed in tables 7 and 8.

For the proportion equations, the results are essentially the same as for the restricted models without Dynarski. The coefficients for RELY are all positive. That they are equal to zero is rejected for the 3RELY model, but not for the LDV model. Given these findings, and the results previously discussed, it seems that economic incentives do appear to influence SAT-taking.

The sum of the RELY coefficients in the score equations in the 3RELY model is positive, and significant. (See Table 9 for test statistics.) However, this result seems to be determined primarily by the estimated coefficient for the non-lagged value of RELY. Only the positive coefficient for $RELY_t$ is statistically different from zero. And again, the coefficient for $RELY_{t-1}$ is negative. In two other tests, the hypothesis that the sum of the coefficients estimated for the lagged values of RELY is zero can't be rejected, nor can the hypothesis that these coefficients are each equal to zero when tested jointly.

The RELY coefficient for the LDV model score equations is positive, but not significantly different from zero at the .05 level. (Equality with zero would be rejected at the less stringent .1 level of significance.)

The results for the LDV model appear to be more in agreement with the hypothesis of relatively slow adjustment

Table 7. Results for 3RELY model, restricted, with values from Dynarski for LAMM and LAMF coefficients.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2606.7 (6.60)	4106.9 (10.3)	2482.1 (6.19)	5342.7 (13.5)	-6.067 (-.85)	-21.730 (3.82)
$RELY_t$	32.29 (2.33)				.21331 (1.82)	
$RELY_{t-1}$	-4.83 (-.33)					
$RELY_{t-2}$	15.46 (1.13)					
YEAR	-1.06 (-5.51)	-1.84 (-9.4)	-1.02 (-5.20)	-2.46 (-12.7)	.00313 (.80)	.01101 (3.96)
ABO	-24.12 (-3.84)	-30.72 (-4.84)	-26.85 (-4.20)	-33.96 (-5.41)	-.29548 (-5.15)	-.22481 (-4.16)
LAMM LAMF	44.82	36.04	44.82	36.04		
D71	3.80 (1.83)				.09702 (5.23)	
MEMB					1.0424 (1.57)	.71237 (1.29)
COST					-1.3271 (-2.10)	
DDRAFT					-1.8194 (-7.13)	
R^2	.9426	.9721	.9506	.9874	.9762	.9594
Durbin-Watson	1.24	2.07	1.39	1.58	2.11	2.15
degrees freedom	32	32	32	32	32	33

Table 8. Results for LDV model, restricted, with values from Dynarski for LAMM and LAMF coefficients.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2332.3 (5.37)	3548.7 (5.68)	2261.0 (4.34)	4386.1 (5.34)	-.79948 (-.09)	-16.693 (-2.60)
$Score_{t-1}$	.20775 (1.90)					
RELY	18.16 (1.56)				.11860 (1.06)	
ABO	-28.05 (-4.76)	-33.42 (-4.84)	-32.00 (-4.97)	-35.88 (-5.13)	-.36638 (-6.06)	-.29362 (-5.18)
LAMM LAMF	44.82	36.04	44.82	36.04		
D71	3.64 (2.31)				.09419 (5.42)	
MEMB					.66007 (.90)	.33388 (.57)
COST					-1.62 (-2.79)	
DDRAFT					1.7935 (6.69)	
YEAR	-.97 (-5.04)	-1.59 (-5.54)	-.94 (-4.01)	-2.02 (-5.23)	.00067 (.17)	.00866 (2.78)
R^2	.9493	.9621	.9633	.9862	.9760	.9656
degrees freedom	29	29	29	29	28	29

of scores. The coefficient estimated for the lagged scores is positive, of reasonable magnitude, and significant as well. Imposing the Dynarski values seems to make the estimates derived from this model correspond more closely to the economic hypothesis.

Table 9. Tests conducted with estimates from restricted 3RELY model, LAMM and LAMF coefficients restricted to values from Dynarski.

Hypothesis	Statistic	Result, .05 level
Coefficients for lagged values of RELY are each zero	F(2,98) = .68	Can't reject
Sum of the RELY coefficients is zero	t(98) = 3.07	Reject

For both models the YEAR and ABO coefficients are significant and indicate strong negative effects on SAT scores. This is true for ABO in the proportion equations also. As before, the estimated coefficient for YEAR is significant and positive in the female proportion equations, while the coefficient for DDRAFT is positive and significant for the males.

The estimated coefficients for COST and MEMB are again of expected sign. And again, only in the case of COST is the null hypothesis of equality to zero rejected.

Given all of the above results, the fully restricted 3RELY model with Dynarski values seems most in agreement with predictions from theory. And an alternative model is suggested by the 3RELY results -- a model in which scores adjust completely within the period in which a change is observed in the return to college education.

The STATIC Model

The relation between the STATIC model and one of the maintained hypotheses should be noted. It is thought that intellectual capital is accumulated over a period of years. For this reason, it was presumed that the response of study habits to changes in incentives will not be fully manifested until several years after the initial change, and that test scores should adjust over time. This hypothesis was brought into the estimation model by means of lagged variables. These hypothesized adjustment patterns did not work well, and the results suggested a model with complete contemporaneous adjustment of scores.

The specification of the model is identical to that of the 3RELY model, with the lagged values of the income variable excluded. The proportion equations are identical to those of the 3RELY model.

The STATIC model is estimated with the full set of restrictions previously described, including the Dynarski values for the LAMM and LAMF coefficients. The procedure used is SUR. The results are reported in Table 10.

Not surprisingly, these results do not differ greatly from those of the 3RELY model. The coefficient for the RELY variable is positive and significant for the score equations, and for the proportion equations (at the .1 level) as well. These estimates are sufficiently large to indicate an important role for RELY in both equations. The

Table 10. Results for STATIC model, restricted, with values from Dynarski for LAMM and LAMF coefficients.

	MMC	VMC	MFC	VFC	CRITM	CRITF
Constant	2879.6 (11.3)	4379.8 (16.1)	2755.0 (9.6)	5615.6 (21.2)	-5.4696 (-.79)	-20.841 (-3.87)
RELY	31.794 (2.69)				.22156 (1.92)	
ABO	-27.87 (-5.28)	-34.47 (-6.22)	-30.60 (-5.30)	-37.71 (-6.93)	-.29323 (-5.21)	-.22354 (-4.22)
LAMM LAMF	44.82	36.04	44.82	36.04		
D71	3.07 (1.79)				.09565 (5.27)	
MEMB					.97282 (1.54)	.61856 (1.20)
COST					-1.3267 (-2.24)	
DDRAFT					1.8388 (7.29)	
YEAR	-1.19 (-9.60)	-1.97 (-14.75)	-1.15 (-8.13)	-2.59 (-20.0)	.00284 (.84)	.01058 (4.02)
R^2	.9510	.9690	.9383	.9864	.9761	.9589
Durbin-Watson	1.25	2.03	1.25	1.50	2.08	2.11
degrees freedom	28	28	28	28	26	27

hypotheses of positive effects for economic incentives on SAT participation and SAT scores stand up well in the light of these results.

There are no important changes for the estimated coefficients on the other explanatory variables, either. As before, the estimates indicate that ABO exerts a strong negative influence in every equation, as does COST in the proportion equations. Trends in omitted variables captured by YEAR have a statistically significant negative effect in the score equations, and a positive effect in the female proportion equation. It can't be rejected that the small negative estimate for the coefficient on YEAR in the male proportion equation is zero, and again the coefficient for DDRAFT is large, positive, and significant. The coefficient for MEMB is positive, but the estimate is not significantly different from zero.

Discussion

The three hypotheses which make up the economic explanation of SAT trends met with differing results. There seems to be evidence for the proposed effect of economic incentives on participation rates, as well as for a direct influence on average scores. On the other hand, the expected effect of increased participation in lowering average scores failed to appear. Indeed, the data suggest just the opposite.

Because the negative effect of increased participation seems to be well documented, the models were estimated with the Dynarski values imposed. These estimates simply assume the selectivity effect and estimate the influence of economic incentives. Imposing these restrictions seems to strengthen the evidence for a role for economic returns in SAT fluctuations.

CHAPTER 6

DISCUSSION AND CONCLUSION

This chapter summarizes the study, with special attention to the empirical results. The meaning of the results for the hypotheses is discussed and the question of exactly how important economic factors may have been in the SAT decline is considered. Finally, several problems with the study are discussed, and suggestions are made for future research.

The objective of this study has been to investigate the hypothesis that fluctuations in the economic returns to education played a significant role in the fall and rise in Scholastic Aptitude Test scores which has occurred since the early sixties. Previous work has indicated that investment in education -- and especially college education -- is closely related to economic returns. An analytical model was thus constructed in which changes in the return to college affect the level of intellectual capital which college-bound high school students acquire, and also affect the proportion of a cohort which chooses to take the SAT. An empirical model was also constructed which consisted of a system of equations determining the proportion of each cohort taking the exam and the average exam scores.

Estimation of the model proceeded from two specifications of the score responses over time to changing incentives, the first using lagged values of relative income (3RELY model) and the second using a lagged dependent variable (LDV model). The score equations were expected to show incomplete contemporaneous adjustment to changes in college returns. Both versions of the model were first estimated with no restrictions imposed, using three stage least squares, and tests were conducted of a variety of possible restrictions on the coefficients. Next, both models were estimated with restrictions imposed on coefficients across exams and sexes. These restricted versions were then re-estimated with the additional restriction of the values derived from Dynarski (1987) for the coefficients of the proportion variables in the score equations. Finally, a simplified model was estimated showing complete adjustment of scores contemporaneously with income changes.

The results of the estimations for the economic hypotheses were mixed. For the 3RELY model, the income variable was found to have the expected positive influence on both proportions and scores. However, the hypothesized adjustment structure for the score equations did not work particularly well. Although the sums of the estimated coefficients for the RELY variables were always positive, in accord with the hypothesis, the structure of adjustment did not make sense. In every case some of the RELY coefficients were negative. However, none of these negative coefficients were statistically different from zero. Also, the hypothesis that the coefficients on the lagged values of RELY were all equal to zero could not be rejected.

These results suggested the STATIC model, which was identical to the 3RELY model except for the absence of lagged income variables in the score equations. RELY was found to have statistically significant and positive coefficients in both score and proportion equations.

The importance of RELY in the LDV model equations was less clear. The estimated coefficients were always positive as expected, but none of the estimates were significant at the .05 level. Thus it was impossible to reject the hypothesis that income has no effect at all.

Estimation of selectivity effects on scores was problematic. Estimates of these coefficients were nearly all negative, of relatively large absolute magnitude, and often significant. This would indicate a strongly positive effect for proportion on SAT scores. Such a finding is opposite to that predicted by the theoretical model, and would seem to cast doubt upon the modeling of the selection effect. But it also must be noted that these results stand squarely against a large body of research on tests and selectivity in general, and on the historical effect of proportion on SAT scores in particular. If the selectivity estimates of the present study are believed, a one percent increase in SAT participation could lead to as much as a 2.3 point gain in average SAT score (using the estimated coefficient from the female math equation, 3RELY, unrestricted estimate). If this is so, the decline in average SAT score overstated any decline in levels of student achievement, and the recovery has overstated any improvement. It must be emphasized that such results are strongly contradicted by previous research, and might be considered perverse.

Alternative estimates for the proportion coefficients were imposed, based upon the aforementioned work of Dynarski (1987). The models thus estimated seemed most in accord with the economic hypotheses, especially the 3RELY model.

What do these results mean for the hypotheses of this study? All estimated coefficients for the RELY variables in the proportion equations were positive. They were statistically different from zero for the 3RELY and STATIC models in all cases, except for males in the unrestricted version of the 3RELY model. The hypothesis that test-taking responds positively to increases in the relative college wage is thus supported by the empirical evidence. It appears that if the return to a college education increases, a larger proportion of a cohort will take the SAT.

It was also hypothesized that the effect of rising returns to education would be higher levels of student investment in the skills which the SAT measures. The estimated coefficients for RELY (or, in the 3RELY model, the sums of the coefficients) were always positive, indicating a positive effect for economic returns on test scores. Thus, income incentives appear to have a positive effect on student achievement.

As discussed above, the hypothesis that economic incentives would also have a negative effect on average SAT scores by inducing more less-able students to participate was rejected. Strong evidence from previous research argued against discarding the hypothesis; it was thus imposed as a maintained hypothesis through coefficients based on Dynarski's work.

A separate issue relating to the STATIC model must also be mentioned. It was suspected that student accumulation of college capital adjusts relatively slowly to changing incentives. This hypothesis was based not upon the theoretical model of Chapter 3, but rather on the assumption that the SAT measures intellectual skills acquired over a period of years. This hypothesis met with somewhat negative results, and the STATIC model seemed to work relatively well. It is interesting to speculate about what this might imply about patterns of accumulation of intellectual capital, and especially for that sort of capital which the SAT tests. It may be that student preparation for the SAT can respond relatively quickly, through the use of practice tests, coaching, and the like.[1] Certainly there is nothing in the present study which sheds light on this issue, but a possible direction for future study is indicated. One potential avenue of inquiry is the effect of changing economic incentives on the demand for SAT preparation, perhaps as measured by consumption of SAT preparation courses and self-study materials.

What do the estimates indicate for the direct effect of college income on SAT scores? According to the estimates for the fully restricted 3RELY model with values from Dynarski imposed, a ten percent increase in the relative return to college induces a 3.24 point change in average SAT score, evaluated at the mean of RELY. (This is the direct effect of RELY on scores and ignores the indirect effect through proportion.) The actual fluctuation of RELY over the period covered by this study was approximately ten percent above and below the mean. This suggests that the direct contribution of declining college income to the score decline might have been approximately 6.5 points on both SAT verbal and SAT math. The actual total declines differed across exam and sex, and the relative importance of the 6.5

---

[1]If so, this might call into question exactly what it is that the SAT measures. Is it a meaningful gauge of intellectual capital acquired through years of effort, or does it simply measure a readily learned ability to take the SAT? There exists a literature which is skeptical of the usefulness of the SAT, and argues for variations on this theme. See for examples Owen (1985) and Crouse and Trusheim (1988).

points accounted for by changes in RELY is given in Table 11.

Table 11. Percent of SAT score decline since 1967 explained by changes in RELY, by exam.

Exam	Total decline (in points)	% explained direct effect of RELY	% explained total effect of RELY
SAT_{mm}	23	28.17	23.74
SAT_{mv}	36	18.00	15.72
SAT_{fm}	27	24.00	20.20
SAT_{fv}	50	12.96	11.31

Of course, RELY affects average scores not only directly, but also indirectly through its effects on the proportion taking the SAT. But if the full effects of changes in relative income on scores as given by the fully restricted 3RELY model with Dynarski's values are considered, relative income becomes only slightly less important as an explanatory variable. The added effect of selectivity is not large. These results are also reported in Table 11.

Changes in relative return to college appear to work fairly well as a partial explanation of the SAT decline, given these results. Especially noteworthy are the implications for math scores, where changes in income may have accounted for nearly a quarter of the drop. The SAT decline is generally thought to have been caused by a variety of factors operating in concert. These results suggest that declining economic incentive was an important factor.

If so, there are important implications for the educational policy debate. The SAT decline is frequently portrayed as having been caused entirely by diminishing student achievement, in turn caused by failures within the educational system or within the culture. The upturn is frequently touted as evidence for the success of policy responses to the decline. The entire debate treats declining SAT scores as a problem to be solved.

The results of this study suggest that in part the SAT decline should not be regarded as a problem at all, but rather the result of a rational response by students to changing economic conditions. So too the rebound in SAT scores appears to be, in part, the response of students to increasing financial rewards to occupations requiring higher education.

It is beyond the scope of this study to determine what drove the fluctuations in relative returns to college. The work of Welch (1979), and Stapleton and Young (1988), suggests market response to changes in labor supply resulting from changes in relative sizes of entering cohorts. Presumably, workers are paid according to the value of their work on the market. For whatever reason, employers changed their evaluations of the relative worth of college and non-college workers. To the degree that these changing labor market conditions influenced the SAT decline, it is difficult to argue that the decline represents a problem. It represents, rather, the effects of students re-allocating their efforts to more productive endeavors, thus enhancing overall economic efficiency.

Of course, these estimates also suggest that seventy five percent or more of the decline is not explained by economic incentives, and this fraction of the drop might represent some failure of educational policy. However, if the effects of increasing average birth order in the cohorts during the SAT decline are considered in conjunction with declining economic returns, as much as forty five percent of the score decline can be explained. (Table 12). Presumably, average birth order is a causative factor outside of the realm of educational policy. This picture of SAT score fluctuations is quite reminiscent of the Stapleton-Young model of educational attainment and cohort size, and Welch's findings concerning the effects of cohort size on returns to education. It might well be argued from the results in Table 12 that the SAT decline was in large part a natural adjustment of patterns of human capital accumulation to the demographic and economic effects of the postwar baby boom.

Table 12. Percent of SAT score decline since 1967 explained by changes in ABO, by exam.

Exam	Total points by ABO	% explained by ABO	% explained ABO and RELY
SAT_{mm}	4.87	46.43	44.93
SAT_{mv}	7.77	37.92	37.30
SAT_{fm}	6.51	44.00	44.30
SAT_{fv}	9.33	30.18	29.96

Certainly a caveat must be offered here. The ABO variable used in the model captures the gamut of demographic effects related to birth order. These may include the negative effects of possible classroom crowding or similar policy related demographic effects. Also, the estimated models in this study did not include any specific policy variables, and the effects of policy changes might well have been captured in the strongly negative and statistically significant coefficients for the time trend variables in the score equations. And there remains at least fifty percent of the decline unaccounted for. The results certainly must be interpreted with caution.

A possible weakness of this study is the use of the income maximization model. It was assumed that the total amount of free time which a student devotes to study and work for financial gain is constant. If this assumption is relaxed, and the value of leisure to the student is considered, changing returns to education might well have an ambiguous effect. For example, if both the college and high school wages increased in such a way that their ratio remained constant, student accumulation of intellectual capital might well decrease, if students were to reduce both study and work in order to enjoy more leisure. It might be that rising U.S. wage rates for all levels of education have encouraged less study. The unavailability of suitable data on the issue of time devoted to leisure made the simpler income maximization approach seem more desirable, but the issue should not be ignored. Perhaps it would prove useful to look at the relation between relative returns to college and student achievement in countries where the high school and college wage differential has been greater and more variable (e.g. South Korea).

One difficulty which plagued this study was the small number of available observations. This, in conjunction with the fact that the dependent variable fell, and then rose, steadily (and was followed in this closely by most of the independent variables), giving only one fluctuation in the time period of the study, made it difficult to obtain precise estimates. It would certainly be desirable to utilize a data set with a larger number of observations, perhaps using a different measure of intellectual achievement. Unfortunately, the College Board's data seem to be the longest running and most thoroughly recorded set available.

A possible way of overcoming this problem would be to utilize cross-sectional data, perhaps after the fashion of Dynarski, examining relations between SAT-taking, returns to education, and other factors between states. Also, if data could be obtained on the background and scores of individual test-takers, better estimates of the selectivity effect might be obtained. Information for individuals concerning parents' levels of education, family income, and high school performance could be particularly useful in this regard.

Another approach which might prove profitable would be to more precisely model the effects of wages on college education, perhaps by analyzing the different types of college major, e.g. engineering or teaching, and looking at the patterns of enrollment in each. Such patterns might well be related to intellectual achievement as well as to changes in relative wages. It is interesting to note that the Student Descriptive Questionnaire given to each SAT-taker asks for information on planned major.

Given the interest that has been focussed on SAT trends, it is remarkable how little attention has been devoted to economic motives as an explanatory factor. This study can offer neither a definitive nor ultimately conclusive account of such, but it certainly points to an important role for economic incentives in the acquisition and measurement of intellectual capital. Further research will be required to more fully assess this role.

REFERENCES CITED

REFERENCES CITED

Becker, Gary S.; Human Capital, First Edition; National Bureau of Economic Research; New York; 1964.

Becker, Gary S.; Human Capital, Second Edition; National Bureau of Economic Research; New York; 1975.

Bishop, John H.; "Is the Test Score Decline Responsible for the Productivity Decline?"; American Economic Review, vol. 79, no. 1, pp.178-197; 1989.

Breland, Hunter M.; "Family Configuration Effects and the Decline in College Admissions Test Scores: A Review of the Zajonc Hypothesis"; College Entrance Examination Board; Princeton, N.J.; 1977.

Congressional Budget Office; Trends in Educational Achievement; Washington , D.C.; 1986.

Congressional Budget Office; Educational Achievement: Explanations and Implications of Recent Trends; Washington, D.C.; 1987.

Crouse, James and Trusheim, Dale; The Case Against the SAT; University of Chicago Press; Chicago; 1988.

Dynarski, Mark; "The Scholastic Aptitude Test: Participation and Performance"; Economics of Education Review, vol. 6, no. 3, pp. 263-273; 1987.

Freeman, Richard B.; "Demand for Education"; in Handbook of Labor Economics, Volume I; Ashenfelter, O. and Layard, R., editors; Elsevier Science Publishers BV; 1986.

Hanushek, Eric A.; "The Trade-off Between Child Quantity and Child Quality: Some Empirical Evidence"; Working Paper no. 97; Rochester Center for Economic Research; University of Rochester; 1987.

Jackson, Rex; "An Examination of Declining Numbers of High-Scoring SAT Candidates"; College Entrance Examination Board; Princeton, N.J.; 1977.

Owen, David; None of the Above: Behind the Myth of Scholastic Aptitude Testing; Houghton Mifflin Company; Boston; 1985.

Schrader, William B.; "Distribution of SAT Scores as an Indicator of Changes in the SAT Candidate Population"; College Entrance Examination Board; Princeton, N.J.; 1977.

REFERENCES CITED, (continued)

Stapleton, David C. and Young, Douglas J.; "Educational Attainment and Cohort Size"; Journal of Labor Economics, vol. 6, no. 3, pp. 330-361; University of Chicago; Chicago; 1988.

Turnbull, William W.; "Student Change, Program Change: Why the SAT Scores Kept Falling"; College Entrance Examination Board; New York; 1985.

Welch, Finis; "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust"; Journal of Political Economy, vol. 87, no. 5, pt 2, pp. S65-S97; University of Chicago; Chicago; 1979.

Wharton, Yvonne L.; "List of Hypotheses Advanced to Explain the SAT Score Decline"; College Entrance Examination Board; Princeton, N.J.; 1977.

Zajonc, R.B.; "Family Configurations and Intelligence"; Science, vol. 192, pp. 227-236; 1976.

Zajonc, R.B. and Bargh, John; "Birth Order, Family Size, and Decline of SAT Scores"; American Psychologist, vol. 35, no. 7, pp. 662-668; 1980.

Zajonc, R.B. and Markus, G.B.; "Birth Order and Intellectual Development"; Psychological Review, no. 82, pp. 74-88; 1975.

APPENDICES

APPENDIX A

TABLE 13: SUMMARY OF DATA

Table 13. Summary of Data

YEAR	SEVM	RELY	DRAFT	TUITPUB	TUITPRIV	MEDY	CEEB	#COLL	SM	SF
1967	1781	.315878	298.6	338	1468	7378	1017	2532	475	385
1968	1849	.338711	341.4	357	1553	7803	1067	2611	491	400
1969	1902	.353174	262.6	377	1638	8619	1097	2731	498	416
1970	1958	.354764	203.7	427	1809	9173	1132	2778	505	430
1971	2009	.333190	153.6	458	1960	9653	1153	2812	461	403
1972	2054	.322928	25.3	490	2111	10771	1168	2867	430	391
1973	2083	.254192	0	521	2261	11671	1190	2932	400	373
1974	2165	.228428	0	552	2412	12381	1167	2992	370	360
1975	2175	.187509	0	599	2614	13207	1155	3022	379	373
1976	2173	.194351	0	643	2861	14094	1145	3026	387	388
1977	2172	.173644	0	688	3107	15331	1149	3086	383	389
1978	2213	.141572	0	732	3354	16574	1153	3095	384	406
1979	2177	.170583	0	777	3487	18174	1157	3134	386	409
1980	2162	.189790	0	840	3811	19974	1136	3152	386	414
1981	2133	.245315	0	915	4275	21022	1132	3231	367	403
1982	2040	.262614	0	1042	4887	21543	1127	3253	372	406
1983	1931	.294402	0	1164	5583	23115	1126	3280	358	390
1984	1878	.263851	0	1270	6140	25489	1130	3284	362	402
1985	1846	.343467	0	1388	6826	25845	1140	3331	373	413
1986	1883	.389189	0	1536	7374	31657	1165	3340	374	423
1987	1930	.374111	0	1651	8118	32821	1166	3406	411	467

Table 13, (continued)

YEAR	JM	JF	HSGM	HSGF	CM	CF	ABO	VMC	VFC	MMC	MFC
1967	263	236	1328	1344	578.7	491.4	2.6058	463	468	514	467
1968	307	282	1338	1357	597.7	510.9	2.6177	464	466	512	470
1969	314	294	1399	1423	607.2	531.2	2.6660	459	466	513	470
1970	311	298	1430	1459	614.6	548.8	2.7068	459	461	509	465
1971	307	302	1454	1483	561.1	515.0	2.7488	454	457	507	466
1972	293	292	1487	1514	523.7	499.0	2.7904	454	452	505	461
1973	281	286	1500	1536	515.9	498.9	2.8189	446	443	502	460
1974	277	279	1512	1561	492.6	492.6	2.8481	447	442	501	459
1975	274	274	1542	1591	496.9	499.6	2.8852	437	431	495	449
1976	279	286	1552	1596	494.6	505.0	2.9191	433	430	497	446
1977	272	282	1548	1607	479.0	500.3	2.9336	431	427	497	445
1978	269	286	1531	1596	478.8	510.4	2.9561	433	425	494	444
1979	275	292	1523	1594	479.2	512.4	2.9590	431	423	493	443
1980	281	300	1491	1552	478.2	513.0	2.9369	428	420	491	443
1981	285	307	1483	1537	478.4	515.6	2.8958	430	418	492	443
1982	291	313	1471	1524	476.3	512.0	2.8226	431	421	493	443
1983	288	309	1437	1451	464.9	498.0	2.6927	430	420	493	445
1984	299	319	1365	1402	464.9	499.8	2.6018	433	420	495	449
1985	307	331	1321	1356	472.0	505.4	2.4870	437	425	499	452
1986	341	371	1303	1339	481.5	519.3	2.4264	437	426	501	451
1987	367	402	1331	1367	520.3	560.1	2.3631	435	425	500	453

Table 13, (continued)

SEVM = SEVENTEEN YEAR OLD MALES
RELY = RELATIVE INCOME
DRAFT = NUMBER OF DRAFT INDUCTIONS
TUITPUB = TUITION PUBLIC - AVERAGE TUITION 4 YEAR PUBLIC INSTITUTIONS
TUITPRIV = TUITION PRIVATE - AVERAGE TUITION 4 YEAR PRIVATE INSTITUTIONS
MEDY = MEDIAN INCOME, MALES EMPLOYED YEAR ROUND, FULL TIME, AGED 45-54
CEEB = MEMBERSHIP ON THE COLLEGE BOARD, NUMBER OF INSTITUTIONS
#COLL = NUMBER OF COLLEGES
SM = NUMBER OF SENIOR MALE TEST TAKERS
SF = NUMBER OF SENIOR FEMALE TEST TAKERS
JM = NUMBER OF JUNIOR MALE TEST TAKERS
JF = NUMBER OF JUNIOR FEMALE TEST TAKERS
HSGM = HIGH SCHOOL GRADUATING SENIOR MALES
HSGF = HIGH SCHOOL GRADUATING SENIOR FEMALES
CM = COHORT MALE TEST TAKERS
CF = COHORT FEMALE TEST TAKERS
ABO = AVERAGE BIRTH ORDER
VMC = SAT SCORE VERBAL MALE COHORT
VFC = SAT SCORE VERBAL FEMALE COHORT
MMC = MATHEMATIC MALE COHORT
MFC = MATHEMATIC FEMALE COHORT

APPENDIX B

GENERATING PROPMC, PROPFC FOR 1967-1971

APPENDIX B

GENERATING PROPMC, PROPFC FOR 1967-1971

The proportion of a cohort taking the SAT is equal to the fraction which takes the exam in the senior year plus the fraction which takes it in the junior year and does not retest in the senior year. (There are students who take the exam in the freshman or sophomore year and do not repeat. This number appears to be quite small and may be ignored.) For males, this cohort proportion may be written as

$$\begin{aligned} PROPMC_t &= PROPMS_t + (1-r)\,PROPMJ_t \\ &= PROPMS_t + PROPMJ_t - r\,(PROPMJ_t) \end{aligned} \qquad \text{(A2.1)}$$

where r is the fraction of junior test-takers which repeats in the senior year.

The fraction r was not constant during the years covered by this study, and appears to have exhibited a strong time trend. In order to capture such a trend, r is specified as a function of time and some constant.

$$r = -(\alpha + \beta t + \gamma t^2) \qquad \text{(A2.2)}$$

Combining (A2.1) and (A2.2), an equation can be derived which will allow estimation of cohort proportion as a function of senior proportion, junior proportion, and time. The resulting coefficients can then be used to generate predicted values of PROPMC and PROPFC for 1967-1971. The resulting specification is

$$PROPMC_t = 1(PROPMS_t + PROPMJ_t) + \alpha(PROPMJ_t) + \beta(PROPMJ_t)t + \gamma(PROPMJ_t)t^2 \qquad (A2.3)$$

This equation is run as ordinary least squares for the years for which cohort numbers are available, 1972-1987. The same procedure is followed for females, using the appropriate values. The regressions give the results found in Table 14.

Table 14. Estimated coefficients for equations used in deriving estimates of PROPMC and PROPFC for 1967-1971.

PROPMJ	(PROPMJ)t	(PROPMJ)t^2
-1984	2.0067	-.000508
(-1.174)	(1.175)	(1.177)
PROPFC	(PROPFC)t	(PROPFC)t^2
-3219.6	3.2628	-.000827
(-2.350)	(2.358)	(-2.366)

(t ratios reported below estimated coefficients, 13 degrees of freedom.)

The coefficients estimated for equation (4.4) can be used to generate predicted values for the missing cohort observations. The values so generated can be used to increase the number of observations, adding degrees of freedom and making estimation of the empirical model a more promising task. These predicted values are reported in Appendix A with the other data.