University of Wisconsin-Madison

Institute for Research on Poverty

Discussion Papers

Robert Meyer

APPLIED VERSUS TRADITIONAL MATHEMATICS: NEW ECONOMETRIC MODELS OF THE CONTRIBUTION OF HIGH SCHOOL COURSES TO MATHEMATICS PROFICIENCY

DP # 966-92

Institute for Research on Poverty
Discussion Paper no. 966-92

**Applied Versus Traditional Mathematics:
New Econometric Models of the Contribution of High
School Courses to Mathematics Proficiency**

Robert Meyer
Department of Economics
La Follette Institute of Public Affairs
Institute for Research on Poverty
University of Wisconsin-Madison

March 1992

This research has been supported by the National Assessment of Vocational Education, U.S. Department of Education, and the Institute for Research on Poverty. The author has benefited from helpful discussions with Charles Benson, Robin Horn, John Wirt, and David Wise and from comments by David Ellwood, Richard Freeman, Art Goldberger, Eric Hanushek, Don Rock, and Aaron Pallas. An earlier draft of this paper was presented at the annual meeting of the American Economics Association.

Abstract

There is no question that traditional mathematics courses such as algebra, geometry, and calculus contribute to the math proficiency of high school students. What is not usually recognized, however, is that math-related and applied courses such as the ones found in science and vocational education programs also appreciably increase math skills. Using a two-stage least squares model and an errors in variables model, the author analyzes the contribution of specific courses to growth in mathematics proficiency during the last two years of high school. Unlike similar previous studies, the econometric models used in the present study control for measurement error in prior achievement--that is, how proficient the students were as sophomores. The author uses his findings to recommend a systemic reform of high school curricula in which math skills are taught in applied contexts and not only as abstract principles.

Applied Versus Traditional Mathematics:
New Econometric Models of the Contribution of High School Courses to Mathematics Proficiency

I. INTRODUCTION

Policy Context

Within the last decade the nation has developed a renewed interest in improving the academic skills of high school students, motivated in large part by widespread concern over the deteriorating performance of high school students throughout the 60's and 70's on the SAT and other national examinations, the exceedingly poor performance of American high school students in international comparisons of math proficiency, and the reported dissatisfaction among employers with the basic skills of young workers (Congressional Budget Office, 1986; Congressional Budget Office, 1987; McKnight et al., 1987).

In mathematics and science this concern has been heightened by the fact that, while most high school students enrolled in the past in a full complement of English courses, many students took no mathematics or science coursework beyond the ninth or tenth grade (see Table 1). Increased enrollments in mathematics and science have therefore emerged as perhaps the most important mechanism for spurring growth in mathematics achievement. In fact, in 1983, the National Commission on Excellence in Education (NCEE) in their report, A Nation at Risk, recommended that states require high school graduates to take a minimum of three years of mathematics and three years of science in grades nine through twelve. The report also recommended that graduates be required to take at least four years of English, three years of social studies, and a 1/2 year of computer science.[1] As indicated in Table 2, less than half of the graduates of the high school class of 1982 would have satisfied the recommended math requirement. Only 30 percent would have satisfied the recommended science requirement.

TABLE 1

The Distribution of Credits Earned in Mathematics and Science: 1982 High School Graduates

Credits Earned	Math	Science
0	0.7%	3.1%
1/2	1.1	2.1
1	11.9	23.1
1 1/2	5.5	5.7
2	27.6	31.3
2 1/2	6.2	4.0
3	21.2	15.5
3 1/2	4.8	2.6
4	16.0	8.5
4 1/2	2.2	1.3
5 or more	2.9	2.6
Mean credits earned	(2.60)	(2.14)

Source: High School and Beyond Sophomore Cohort Transcript Survey.

Note: Percentages do not add to 100.0 due to rounding.

TABLE 2

The Share of Students Satisfying the NCEE's Minimum Graduation Requirements: 1982 High School Graduates

Minimum Graduation Requirements	Share Satisfying Requirements
3 or more math credits	47.0%
3 or more science credits	30.4
3 or more math and 3 or more science credits	24.7
4 or more English credits	58.3
3 or more social studies credits	65.2
4 or more English credits and 3 or more social studies credits	42.9
3 or more math credits, and 3 or more science credits, and 4 or more English credits, and 3 or more social studies credits	14.1

Source: High School and Beyond Sophomore Cohort Transcript Survey.

Note: NCEE = National Commission on Excellence in Education.

State governments responded rapidly to the demands for tightened graduation requirements. In fact, within twelve months of the publication of A Nation at Risk thirty-five states introduced or strengthened their graduation requirements (Bennett, 1988). Their response, however, fell far short of the Nation at Risk recommendations. Only ten states adopted the recommended math requirement of three years and only three states adopted the recommended science requirement of three years (Meyer, 1990). In contrast, thirty-nine states implemented (and two states recommended to local districts) the four-year requirement in English.[2]

Given the failure of most states to implement the ambitious math and science reforms articulated in A Nation at Risk, what can be done to promote development of mathematics skills in secondary schools? To answer this question, it is necessary to identify the factors that do, in fact, contribute to mathematics learning. In light of the current focus on course-taking requirements, this paper concentrates on identifying the contribution of secondary school courses to mathematics proficiency. The major new finding is that mathematics learning is substantial in many courses other than traditional mathematics courses, particularly for the most academically disadvantaged students, the group that takes the least amount of traditional mathematics. Moreover, many "math-related" courses are in subjects such as vocational education that are quite popular with students, particularly in contrast with math courses. These empirical findings indicate that learning mathematics in an applied context is a viable alternative or complement to enrolling in traditional mathematics. In the concluding section of this paper I draw on these findings to articulate a new curriculum-wide, systemic strategy for spurring growth in mathematics proficiency. This approach rejects the traditional view as exemplified in A Nation at Risk, that mathematics instruction is the sole province and responsibility of high school mathematics departments. The evidence suggests that given existing courses (that is, the current "technology of production"), a systemic approach to mathematics learning--in particular, shared responsibility for math instruction across vocational education, science, and traditional mathematics--could more than triple the mathematics learning of academically disadvantaged high school students. Presumably, new vocational and applied science courses, explicitly designed with a mathematics focus, could generate even more impressive gains. The idea of integrated academic and vocational education is not, of course, without precedent. Writing-across-the-curriculum programs, which require that writing skills be taught in all courses, not just English, reflect a similar motivation.

Methodological Issues

In developing a curricular model of mathematics learning, this study gives careful consideration to the econometric problems inherent in estimating the causal (value-added) contribution of courses to mathematics proficiency. Since course enrollment choices are heavily influenced by prior mathematics achievement, I find that estimates of the effects of courses on growth in mathematics achievement are extremely sensitive to model misspecifications. In particular, models that control for prior achievement using proxies (e.g., family background) rather than actual prior achievement yield severely biased parameter estimates. In addition, measured prior achievement itself is a very imperfect control for prior achievement due to unavoidable error in its measurement.

I offer two alternative models for obtaining consistent parameter estimates in the presence of measurement error in prior achievement: an instrumental variables/two-stage least squares estimator (2SLS) and an errors in variables (EV) estimator. Both models maintain the assumption that the unexplained components of achievement growth are uncorrelated from year to year. I present some evidence that this may not be a bad assumption. A more general investigation of this issue awaits the availability of data containing at least three periods of achievement data.

The first method, which has not previously been used in educational production function research, is motivated by a recursive structural model of curricular enrollments and mathematics proficiency. I show that the math proficiency equation can be estimated using two-stage least squares

(2SLS). The key instruments in the first-stage equation for prior math achievement are prior course enrollments. The second method uses an external estimate of the variance of the measurement error in prior achievement to obtain consistent parameter estimates. Fortuitously, both approaches yield remarkably similar parameter estimates, which suggests that my empirical results are quite robust. These methods may have important applications in other models of educational production.

Plan of the Paper

Section II articulates a curricular model of mathematics learning that relates growth in mathematics proficiency from the end of tenth to the end of twelfth grade to courses taken during that two-year period. The section begins with a brief review of related literature, in particular, studies of educational production functions and public and private school effectiveness.

Section III describes the course enrollment data used in our empirical analysis. The data were taken from the base year (spring 1980), first follow-up (spring 1982), and high school transcript surveys of the High School and Beyond/Sophomore Cohort (HS&B) study. This survey is the most up-to-date source of information on high school enrollments and test score performance (measured prior to and after participating in courses).

Section IV assesses the statistical models that have been used in previous analyses of the determinants of academic achievement. This literature has generally failed to adequately address the problem of measurement error in tests. As mentioned above, I propose new two-stage least squares and errors in variables models of the determinants of mathematics learning that are derived from a structural model of secondary enrollment choices and mathematics learning.

Section V compares the performance of alternative models of mathematics learning.

Section VI presents estimates for college-bound students and academically disadvantaged, nondisadvantaged, and advantaged students, using the preferred estimators identified in the previous section. Among the preferred estimators, the statistical results are exceptionally robust. Alternative 2SLS and errors in variables models yield very similar parameter estimates.

Section VII concludes the paper with an analysis of the implications of our empirical analysis for specific academic reforms.

II. A CURRICULAR MODEL OF MATHEMATICS LEARNING

Review of the Literature

Since the publication of the "Coleman Report" (Coleman et al., 1966) just over two decades ago, numerous researchers have examined the links between academic achievement, as measured by student performance on standardized tests, and school quality. At face value, their research findings seem surprising and even contradictory. On the one hand, studies by Hanushek (1971), Murnane (1975), and others have found, as expected, that some teachers and schools contribute substantially to measured student achievement. On the other hand, obvious measures of school and teacher quality and school inputs, such as per-pupil expenditures, student/teacher ratios, and teacher experience, were found not to be strongly or systematically related to student performance. Instead, the primary influences on student performance were student and family backgrounds--characteristics that cannot easily be changed in the short run (if at all) to improve student achievement (Hanushek, 1986).[3]

Although these findings may be due, in part, to poor measurement of school inputs, an alternative explanation is that student and school performance is determined primarily by factors other than conventionally defined school inputs, for example, the organization and process of education, the curricular content of education, student effort, and the quality of school inputs within the context of specific student and classroom needs.[4] Recent research, using analytic methods similar to those developed in the school input/educational production function literature, has begun to explore the importance of these factors.

In particular, the controversial research by Coleman, Hoffer, and Kilgore (1982) has stimulated a lively debate over the authors' conclusion that private schools are better than public schools in promoting academic achievement.[5] The authors' original research was controversial, in part, because it was based on cross-sectional data from the base-year survey of the High School and Beyond (HS&B) study, rather than on longitudinal data containing test score data for the same cohort over time. Recently, Hoffer, Greeley, and Coleman (1985) and Alexander and Pallas (1985) used base-year and follow-up data for the HS&B sophomore cohort to examine the contribution of public and private schools to the change in student test scores from the tenth to the twelfth grades. (Our empirical analysis is also based on these data.) Despite the fact that the two papers relied on very different statistical methods,[6] they both concluded that private high schools have an advantage over public high schools in developing academic skills that is roughly equivalent to two-thirds of a year's (Alexander and Pallas) to one full year's (Hoffer et al.) growth.[7] An intriguing aspect of the Hoffer, Greeley, and Coleman paper of particular relevance here is the finding that higher average enrollments in advanced mathematics courses for private school students fully account for the private school advantage in mathematics achievement growth.[8] This, of course, implies that advanced math courses contribute substantially to the development of mathematics skills.

Researchers who have examined this question have concluded, almost without exception, that enrollment in advanced mathematics is a powerful determinant of mathematics proficiency.[9] Welch, Anderson, and Harris (1982), drawing on data from the seventeen-year-old wave of the 1977-1978 National Assessment of Educational Progress (NAEP) in mathematics, found, as in previous educational production function studies, that family and community characteristics accounted for a large share (25 percent) of the variance in NAEP mathematics achievement. Total enrollment in mathematics courses (as reported by the students), however, explained an additional 34 percent of the variance in mathematics achievement, raising the total variance explained to 59 percent.[10] Schmidt

(1983) reported similar results in his analysis of the National Longitudinal Study of the High School Class of 1972: personal characteristics and hours of instruction in mathematics and other subjects in grades ten through twelve explained 57 percent of the variance in mathematics achievement. Although these studies, as we shall see, were correct in identifying high school mathematics coursework as perhaps the major determinant of mathematics-skills development while in high school, they undoubtedly overstated its contribution to the level of math proficiency, because they did not (and could not, given the data) control for prior (to high school) mathematics achievement.

Pallas and Alexander (1983), on the other hand, analyzed a data set, the ETS Study of Academic Prediction and Growth, that contained measures of math achievement as of the middle of the twelfth grade and the beginning of ninth grade, as well as detailed high school transcripts. Unfortunately, different tests were administered to the students in these two years, the School and College Ability Test (SCAT) to ninth graders and the Scholastic Aptitude Test (SAT)--or its psychometric equivalent, the Preliminary Scholastic Aptitude Test (PSAT)--to twelfth graders. The authors felt that this would not corrupt their findings, arguing that the quantitative subtest of the SCAT (SCAT-Q) was a reasonable proxy for the math component of the SAT (SAT-M). Indeed, they reported that personal and family characteristics plus the SCAT-Q accounted for 57 percent of the variation in the senior-year SAT-M score, with course enrollments explaining an additional 12 percent. The latter percentage is substantially less than that in either the Welch, Anderson, and Harris (1982) or Schmidt (1983) studies, thus confirming the need to adequately control for prior achievement in estimating the value added by courses or other school inputs.

Pallas and Alexander (1983), unlike the papers cited earlier but as the present one will, estimated the contribution of individual courses, rather than total years or semesters of mathematics coursework, to growth in mathematics proficiency. Course variables were derived from high school transcripts and thus were likely to be relatively free of error.[11] They included thirteen different mathematics courses in their analysis, as well as three math-related courses: physics, quantitative business, and quantitative industrial arts (i.e., drafting and drawing). At face value, their estimates indicate that growth in mathematics proficiency is spurred exclusively by advanced math or math-intensive courses such as geometry, trigonometry, calculus, and physics; algebra 1 and 2 make no contribution to the development of mathematics skills, and the contributions of general math 1, applied math, and quantitative business are actually negative, although generally not statistically significant.

In the value-added context, the negative coefficient estimates are somewhat implausible since they indicate that students participating in such courses experienced an attendant loss of mathematics skills. In fact, the heavy tilt of the estimates--large positive coefficients for advanced courses, zero coefficients for intermediate courses (algebra 1 and 2), and negative coefficients for low-level courses--strongly suggests that the course coefficients reflect something other than the value added or causal effect of the courses, such as perhaps the fact that the ninth grade SCAT-Q test is an imperfect control for the twelfth grade math SAT test, the authors' outcome variable.

But a possibly more important factor is that Pallas and Alexander (1983) failed to control for the fact that the SCAT-Q, like all tests of finite length, is subject to measurement error. Failure to control for measurement error, of course, is not unique to the Pallas and Alexander (1983) paper, but its consequences are likely to be particularly severe in a model that includes detailed course variables.[12] These variables tend to be highly correlated with prior test scores, thereby grossly magnifying the consequences of even modest measurement errors in test scores. The evidence presented in Section V suggests that the coefficient estimates in Pallas and Alexander (1983) are, in fact, severely biased due to measurement error in the SCAT-Q test.

Study Objectives and Definition of Course Variables

The objective of this study is to estimate the extent to which high school mathematics and math-related courses contribute or have the potential to contribute to the growth of mathematics proficiency. In order to estimate the contribution to mathematics proficiency of alternative math and math-related courses, I estimate a "curricular" educational production function. The exact specification of this model is affected by the following two considerations, both of which are unique to a model of curricular effectiveness. First, most high school students enroll in five to six courses at any point in time. Since test batteries were administered to the students in the HS&B survey during the spring of tenth and twelfth grade, the students could have taken eighteen or more courses (some lasting a single semester) during the intervening period. In essence, students were exposed to multiple rather than single treatments, which means that even simple estimates of course effectiveness must be derived using multiple regression methods. Alternatively, it might be possible to categorize students into mutually exclusive programs or course enrollment patterns, and then estimate the effectiveness of alternative patterns.[13] This approach could be attractive, for example, in evaluating the relative effectiveness of new demonstration programs. The disadvantage of this approach, however, is that it fails to identify the factors that lie behind program effectiveness. Since the purpose of this paper is to identify the effectiveness of particular courses, the model needs to include a fairly detailed list of courses.

In most of the models estimated in this study, I include course variables without allowing for possible interactions. In future work, it would be interesting to investigate the degree of substitutability and complementarity among different courses. Unfortunately, the data used in this study (High School and Beyond) are simply too weak to support a more elaborate specification. The main problem is that measurement error in the growth in mathematics proficiency (the outcome variable) is enormous, approximately 70 percent of the variance of the change in math test scores (see

Table A-1). This error variance is substantially larger than the error variance of either the pre (tenth grade) or post (twelfth grade) math test scores, due to the fact that the true change in proficiency from tenth to twelfth grade varies substantially less than the level of math proficiency. In addition, measurement error in the change in math scores is the sum of measurement error in the sophomore- and senior-year tests, given the assumption of uncorrelated measurement errors in the two tests, a reasonable assumption. The large error variance in the change in mathematics proficiency is essentially equivalent to a reduction in sample size, relative to tests without error, of over 75 percent.[14] As a consequence, the standard errors in the models estimated in this paper are surprisingly large, given that the sample contains well over 10,000 students, thereby limiting the opportunity to explore more flexible functional forms.

The second consideration is the extent to which it is permissible to aggregate courses into aggregated field- or subject-level variables in order to limit the number of explanatory variables in the model. Some aggregation is clearly necessary since the transcripts in the HS&B data base contain over one thousand different course titles, from approximately one thousand high schools. The consequences of course aggregation are illustrated by considering the effect of aggregating two course-enrollment variables, say X_{1i} = general math and X_{2i} = trigonometry. Then, the contribution of X_1 and X_2 is given by

$$\alpha_1 X_{1i} + \alpha_2 X_{2i} + \alpha_{12}(X_{1i} + X_{2i}) + (\alpha_1 - \alpha_{12})X_{1i} + (\alpha_2 - \alpha_{12})X_{2i},$$

where $(X_{1i} + X_{2i})$ is the new course aggregate and α_{12} is the coefficient on the course aggregate. The two terms on the far right become part of the residual error. If α_1 and α_2 are equal, α_{12} will equal α_1 and α_2 and the two terms on the far right drop from the equation. This well-known result indicates that it is permissible to aggregate variables that share the same coefficient. Of course, a priori, it is difficult to know which course variables share the same coefficients, except, perhaps, for the large number of courses that are hypothesized to have no effect on mathematics learning, for example, physical education, health, social studies, foreign languages, and English.

Aggregation of courses with dissimilar coefficients will lower the explanatory power of the model by adding to the variance of the error term. If the variation in coefficients is relatively small, however, aggregation will yield a more precise estimate of the (weighted) average effect of the course aggregate. This average coefficient is similar to the mean coefficient in a model with random coefficients. Unfortunately, this procedure may also create severe bias in many, if not all, of the parameters in the model, including those for which aggregation is perfectly permissible. In the example discussed above, this would occur if (1) enrollment in general math (X_{1i}) is negatively correlated with a course variable such as foreign language enrollment, (2) enrollment in trigonometry (X_{2i}) is positively correlated with foreign language enrollment, and (3) the trigonometry coefficient α_2 is greater than the general math coefficient α_1. Under these plausible conditions, the foreign language coefficient will be biased upward because it acts as a proxy for enrolling in trigonometry rather than general math. In order to minimize this problem, I refrain (where feasible) from aggregating courses that are hypothesized to have large, but possibly different, effects on mathematics learning. In particular, I include eleven different mathematics variables in the analysis, ranging from general math to algebra 1 to calculus. In total, the model includes nineteen different course variables. Previous analyses have typically included fewer course variables. As reported by Koretz (1988), these studies have tended to generate some peculiar results, such as a large apparent effect of foreign language instruction on mathematics learning. Improper aggregation of courses may be one reason for these anomalous results. Some studies, incidentally, also exclude important course variables from their analysis, such as science coursework, thereby introducing old-fashioned omitted variable bias.

As mentioned above, the curricular model of mathematics learning estimated in this study distinguishes nineteen different course-, field-, or subject-level variables, fourteen of which are hypothesized to have large positive effects on mathematics learning. The latter variables include both traditional mathematics courses, applied/vocational math courses, and math-related courses drawn from the sciences and vocational education. Although other subjects may contain math-related courses, in this study, I focus only on math-related courses in science and vocational education.

The premise that high school vocational or science courses can and do contribute to the development of mathematics (or, more generally, academic) skills is based on the notion that the applied, often "hands-on," orientation of these (and other) courses stimulates student interest in learning and provides concrete opportunities to learn the abstract principles taught in mathematics, English, and other "core" subjects. To the extent that individual learning styles differ, some students may actually learn mathematics more readily in the applied context offered by these courses than in the formal setting typically present in traditional mathematics courses. Thus, applied courses could complement or even substitute for traditional math courses in the production of mathematics skills.

By design, the course variables included in the analysis span a continuum that ranges from formal, abstract mathematics to applied mathematics. In addition, some courses involve full-time study of mathematics, while others involve part-time, perhaps incidental, study of mathematics, for example, the math-related science and vocational courses. I emphasize these distinctions because I believe that my empirical evidence can be assessed at two levels. First, the evidence can be interpreted as estimates of the average effectiveness of particular courses in producing mathematics skills, as of 1980 to 1982. To the extent that curriculum offerings now and in 1982 are reasonably similar, the model estimates can be used to predict the current consequences of alternative course-enrollment patterns, for example, increased enrollments in algebra 1. In other words, these estimates can be interpreted as the current "technology" for producing mathematics skills.

Second, the evidence can be used to test the general proposition that mathematics can be learned efficiently in an applied context--for example, in a science lab or in a vocational workshop.

Unfortunately, the HS&B data offer only a weak test of the potential of applied courses to promote math-skills development because the math-related courses identified in the analysis were probably not designed with this goal in mind. Nonetheless, strong positive evidence that applied courses promote mathematics proficiency would establish the principle that mathematics can be learned outside of traditional mathematics courses. Such a finding would suggest ways of improving the technology of mathematics production, perhaps by vigorously expanding applied math and math-related courses and/or by linking together traditional mathematics and math-related courses. This issue is explored more fully later in the paper in light of the empirical evidence.

The formal, abstract mathematics courses included in the analysis are, in ascending order of difficulty, basic math, general math, computer math, prealgebra, algebra 1, geometry, algebra 2, precalculus (algebra 3, trigonometry, advanced geometry, and mathematic analysis), and calculus. These course variables are reasonably well defined and reflect minimal or no amounts of aggregation. Enrollments in these and other courses are discussed in Section III.

Two applied math variables, each with a full-time focus on mathematics instruction, are included: applied math and specific vocational math. The course titles included in these variables are listed below (with Classification of Secondary School Course [CSSC] codes in parentheses):

Applied Math

- consumer mathematics (27.0114)
- mathematics as a liberal art (27.0109)
- mathematics for employment (27.0110)
- science mathematics (27.0108)
- technical mathematics (27.0111)
- vocational mathematics (27.0110)
- other applied mathematics (27.0300)

Specific Vocational Math

- agricultural mathematics (01.0151)
- business mathematics 1 (07.0171)
- business mathematics 2 (07.0172)
- nurse's mathematics (17.0651)

The essential difference between the two variables is that specific vocational math courses are structured around a particular vocational subject such as business and presumably draw most or all of their examples from this subject. Specific vocational math courses may also tend to be taught by vocational educators rather than math teachers. The HS&B data, however, provide no information on this point.

Vocational education courses were split into two groups, math-related and non-math-related.[15] Courses were designated as math-related or non-math-related on the basis of course titles and descriptions. Although these designations were subjective and therefore subject to error, they were influenced by a well-defined set of guidelines. A vocational course was considered math-related if:

- It was one of several vocational/science specialties, such as agricultural science;
- It was any of the following course types:
  - architecture/drafting,
  - accounting/bookkeeping,
  - computer programming,
  - data processing,
  - economics/finance/investments/taxation,
  - electricity,
  - electronics,
  - insurance,
  - real estate;
- It relied on geometric or spacial skills, for example, drafting and graphics courses.

The courses included in the math-related group are listed in Appendix B. These course classifications were made prior to any data analysis. Thus, the empirical results are not contaminated by "data mining." To the extent that some vocational courses have been misclassified in one category or another, the estimated differential between math-related and non-math-related courses is apt to be understated. (In practice, however, the estimated difference between the two groups of courses is substantial, as discussed below.)

Finally, science courses were separated into two groups: chemistry and physics, and biology and survey science. The former group was hypothesized to be more math-related than biology or survey science.

Summary

In this section, I have defined the curriculum variables that are included in my curricular model of mathematics learning. In the next section, I discuss the data used in the empirical analysis.

III. DATA

High School and Beyond Study

The data used in this study were derived from the sophomore cohort of the High School and Beyond study, a nationally representative sample of tenth grade students as of 1980. Participants in the study were surveyed and tested during the spring of 1980 and again two years later.[16] The particular sample used in the research consists of 10,961 students who had valid test-score data and complete transcript information for all four years of high school. This sample includes 10,106 high school graduates and 855 students who dropped out of high school after participating in the spring 1980 (base-year) HS&B survey. Courses in the transcript file were originally coded according to the Classification of Secondary School Courses (CSSC), a classification containing well over a thousand defined and described high school courses. These courses, in turn, were grouped into fields and subject areas using the Secondary School Course Taxonomy (SST), presented and discussed in the First Interim Report of the National Assessment of Vocational Education (1988). (See also Brown et al., 1989.) The distinction between math-related and non-math-related vocational education was made following the rules given in Section II. All tables report course enrollments in terms of standard Carnegie credits.[17] Course enrollments reflect all passed and failed courses.[18]

A battery of six tests was administered to the sophomore HS&B sample during the spring of their sophomore year and again exactly two years later. The tests covered mathematics, the primary test used in our analysis, as well as reading, vocabulary, writing, science, and civics. Although the tests, at face value, measure competencies in six distinct areas, Rock et al. (1985) conducted a factor analysis of the tests that indicated that two underlying factors--a math and a verbal factor--account for essentially all of the legitimate (error-free) variance in the six test scores. In addition, the estimated reliability of the civics test was found to be exceptionally low (approximately 50 percent--see Appendix Table A-1). As a result, the empirical analysis is based on a subset of the tests: the mathematics test, a composite verbal score (the sum of reading and vocabulary test scores), and the science test. The verbal and science tests are included as additional explanatory variables in the econometric model of mathematics learning, to pick up the possible effect on mathematics learning of skills other than those measured by the HS&B math test.[19] Total testing time was twenty-one minutes for the mathematics test and forty-seven minutes for the five other tests. The math test consisted of thirty-eight items, eighteen involving arithmetic skills, twelve involving algebra skills, and eight involving geometry skills. The inclusion of test items related to specific high school mathematics courses--algebra 1 and geometry--raises the possibility that the HS&B math test taps both cognitive mathematics ability and specific knowledge of algebra and geometry skills. If so, estimates of the contribution of algebra 1, geometry, and algebra 2 to the growth of mathematics skills may be overstated, relative to other courses such as prealgebra, trigonometry, calculus, applied math, and math-related vocational education.

In order to allow for the possibility that the math effectiveness of different courses varies in the population, I typically disaggregate the data into college-bound and non-college-bound students and low-, mid-, and high-math-proficiency students. As discussed more fully in Section VI, students were classified into math proficiency triptiles (thirds) on the basis of predicted rather than actual sophomore math achievement.[20] Approximately 60 percent of the sample identified themselves during the base-year (sophomore-year) survey as students who expected to attend college. The remaining 40 percent indicated no plans to obtain postsecondary education (25 percent) or plans to obtain postsecondary vocational training (15 percent).

In the remainder of this section I present estimates of course enrollment in the key course and subject areas described in Section II and descriptive statistics for the sophomore- and senior-year mathematics tests.

High School Course Enrollments

As reported in Table 3, students in the HS&B sample took an average of 2.72 credits in vocational education and applied/specific vocational math during their junior and senior years, approximately 27 percent of all credits. As expected, vocational education (including applied/specific vocational math) accounted for a larger share of the total credits of non-college-bound graduates (38 percent) than for either college-bound graduates (20 percent) or high school dropouts (29 percent). Of course, individuals who dropped out after the spring of their sophomore year took predictably few credits, on average, in all subject areas. Vocational courses identified as math-related accounted for only one-fifth of all vocational credits, slightly more than a half-credit, one-semester course, on average. Applied/specific vocational math represented only 4 percent of all vocational courses. Coursework in math-related vocational education was nearly identical for college-bound and non-college-bound students.

College-bound and non-college-bound students differed sharply in terms of their average coursework in mathematics, science, and foreign languages. College-bound graduates earned 2 1/2 times as many credits in mathematics and science as non-college-bound graduates and more than four times as much foreign language instruction. In fact, non-college-bound graduates took only slightly more mathematics and science than dropouts. College-bound graduates also tended to take more

TABLE 3
Average Course Enrollments in Eleventh and Twelfth Grade Courses by Graduation Status and Post-High School Plans

Course	High School Dropouts	High School Graduates/ Non-College-Bound	High School Graduates/ College-Bound	All Students
Vocational education				
Math-related voc.	0.157	0.576	0.505	0.509
Non-math-related	1.056	3.143	1.503	2.110
All vocational courses	1.213	3.719	2.008	2.619
Specific voc. math	0.041	0.053	0.030	0.040
Applied math	0.051	0.088	0.043	0.061
Mathematics				
Basic	0.040	0.032	0.009	0.020
General	0.104	0.108	0.064	0.084
Computer math	0.007	0.014	0.036	0.025
Prealgebra	0.016	0.022	0.027	0.024
Algebra 1	0.035	0.060	0.077	0.067
Geometry	0.048	0.089	0.186	0.139
Algebra 2	0.038	0.094	0.412	0.263
Precalculus	0.013	0.047	0.316	0.191
Calculus	0.000	0.005	0.089	0.050
All math courses[a]	0.301	0.471	1.216	0.863
Science				
Survey	0.108	0.096	0.063	0.079
Biology	0.126	0.163	0.277	0.223
Chemistry	0.020	0.092	0.475	0.296
Physics	0.011	0.035	0.260	0.155
All science courses	0.265	0.386	1.075	0.753
English	0.928	1.805	1.956	1.828
Social studies	0.862	1.808	1.893	1.790
Fine arts	0.256	0.644	0.718	0.658
Foreign languages	0.062	0.130	0.559	0.359
Personal and other	0.595	1.036	1.164	1.076
Total credits	4.574	10.140	10.662	10.047
Sample size	855	3,759	6,347	10,961

Source: Author's computations from the high school transcripts collected as part of the High School and Beyond study.

Note: Course enrollments are measured in terms of standard Carnegie credits. A one Carnegie credit course typically meets for five fifty to fifty-five minute periods per week for an entire school year. A typical one-semester course would earn one-half credit.

[a]Excluding specific voc. math and applied math courses.

advanced mathematics courses than non-college-bound students and substantially more chemistry and physics.

These numbers suggest that from the limited perspective of mathematics-skills development, the effectiveness of math-related and non-math-related vocational education is especially important for non-college-bound students. During their junior and senior years the non-college-bound graduates took roughly equivalent amounts of mathematics and math-related vocational education (about one semester in each), but five to six times as much non-math-related vocational education (almost six semesters).

Mathematics Test Scores

Table 4 summarizes the performance of students on the HS&B mathematics test. The average sophomore score in 1980 was 13.44, with a standard deviation of 9.70. The average gain in test scores from the tenth to the twelfth grade was 1.92 points, approximately 1 point per year. Although this change was modest relative to the (accumulated) variation in tenth grade achievement, as argued by Jencks (1985) and Hoffer, Greeley, and Coleman (1985), this does not necessarily imply that the observed gain in math scores was small and inconsequential, but rather that individual variation in growth in mathematics achievement from preschool through tenth grade was substantial. Table 4 indicates that high school dropouts scored poorly on the sophomore test (5.78) and failed to score appreciably better two years later. In contrast, non-college-bound and college-bound graduates had higher sophomore scores and increased their test scores by 1.08 and 2.71 points, respectively, from the tenth to the twelfth grade.

Variation in mathematics learning, as will be demonstrated, depends critically on the number and types of courses taken by high school students. However, before discussing my best estimates of the contributions of different courses to mathematics development, it is informative to examine Table 5, which reports the average gain in math scores for students with different levels of total

TABLE 4

Average Math Test Scores by Graduation Status and Post-High School Plans

	High School Dropouts	High School Graduates/ Non-College-Bound	High School Graduates/ College-Bound	All Students
Sophomore score	5.78	9.46	17.22	13.44
	(6.86)	(8.10)	(9.35)	(9.70)
Senior score	6.15	10.54	19.93	15.35
	(7.50)	(8.61)	(10.02)	(10.65)
Gain in score	0.37	1.08	2.71	1.92
	(5.78)	(5.92)	(5.72)	(5.87)
Sample size	855	3,759	6,347	10,961

Source: Author's computations from the high school transcripts collected as part of the High School and Beyond study.

Note: Standard deviations are reported in parentheses.

TABLE 5

Average Sophomore Math Test Scores and Gain in Senior Test Score, by Number of Math Credits in Eleventh and Twelfth Grade and Post-High School Plans

	Non-College-Bound		College-Bound		All Students	
Math Credits	Soph. Score	Gain Score	Soph. Score	Gain Score	Soph. Score	Gain Score
0	8.50	0.13	12.43	0.41	9.75	0.22
1/2	8.37	1.15	12.93	1.64	10.21	1.35
1	9.16	1.58	15.61	2.23	12.99	1.96
1 1/2	9.30	1.58	17.54	2.94	15.45	2.60
2	10.87	3.03	20.68	4.19	18.88	3.98
2 1/2 or more	12.78	3.59	21.91	4.37	20.77	4.28
Overall mean	8.94	0.97	17.01	2.67	13.44	1.92
Sample size	4,448		6,513		10,961	

Source: Author's computations from the high school transcripts collected as part of the High School and Beyond study.

mathematics enrollment over four years. These results cannot be used to infer the contribution of mathematics courses to mathematics learning, because they ignore the influences of other math-related courses and possible variation in the effects of different math courses. Nonetheless, Table 5 suggests that mathematics instruction has a powerful effect on the development of mathematics skills. Students who took no mathematics during their junior/senior years essentially failed to improve their math scores from the tenth to the twelfth grade. In contrast, students who enrolled in even a half credit significantly improved their mathematics proficiency, and students with the highest math enrollments scored the greatest gains. College-bound students improved their scores somewhat more than non-college-bound students, perhaps due to the fact that college-bound students tended to enroll in more advanced courses than non-college-bound students (see Table 3).

IV. ECONOMETRIC METHODOLOGY

The primary objective of this section is to develop models of mathematics proficiency which can be estimated consistently despite the presence of measurement error in prior achievement. To provide the proper context for this discussion, I begin by describing a system of enrollment and learning equations. This provides a framework for describing the parameter bias that afflicts many statistical models commonly used to study the determinants of educational outcomes. In Section V, I estimate and compare the models discussed in this section in order to validate and determine the empirical importance of the theoretical arguments presented below. To simplify presentation, student subscripts (i) are generally omitted in the equations that follow.

A System of Enrollment and Learning Equations

The equation system presented below consists of five equations: equations for true math proficiency (measured without error) at the end of periods 0, 1, and 2 (T_0, T_1, and T_2, respectively) and course enrollment equations for periods 1 and 2 (X_1 and X_2, respectively). The structure of the model is essentially recursive. Enrollments are assumed to depend on lagged (prior) achievement and personal characteristics (Z), and achievement is assumed to depend on lagged enrollments, lagged achievements, and personal characteristics.

(4.1) $T_0 = Z\delta_0 + e_0$

(4.2) $X_1 = ZD_1 + T_0b_1 + w_1$

(4.3) $T_1 = Z\delta_1 + X_1\alpha_1 + \theta_1T_0 + e_1$

(4.4) $X_2 = ZD_2 + X_1C_2 + T_1b_2 + w_2$

(4.5) $T_2 = Z\delta_2 + X_2\alpha_2 + \theta_2T_1 + e_2$

Since X_1 and X_2 are vectors, C_2, D_1 and D_2 represent parameter matrices, and b_1 and b_2 are parameter vectors. In our context, periods 0, 1, and 2 correspond to birth through eighth grade, ninth and tenth grade, and eleventh and twelfth grade, respectively. The system can obviously be extended to more than three periods. As is, equation (4.1) depends only on personal and family characteristics (Z) and thus should be interpreted as a reduced-form equation in Z.[21]

In the absence of measurement error in test scores T_0 and T_1, ordinary least squares could be used to obtain unbiased estimates of all five equations under the following conditions: (1) the enrollment equation errors (w) and the math proficiency equation errors (e) are uncorrelated and (2) the math proficiency errors (e) are not correlated over time.

The first condition is violated if the effects of courses vary randomly across individuals and students know of and are responsive to that variation. This is a possible source of selection bias that is not explicitly captured in this study. The first condition is also violated if particular enrollment choices (say, enrollment in calculus) tend to be correlated with unobserved experiences that produce math gains (for example, participation in math clubs). Exclusion of such experiences, of course, could cause omitted variable bias in estimates of course effectiveness (α). However, given that most mathematics learning probably occurs in classrooms rather than in unobserved extracurricular activities, this is probably not a serious problem. This issue is likely to be more serious in models of verbal achievement, however.

The second concern, correlation over time in the achievement equation errors (e_0, e_1, and e_2), is potentially important. For example, students with high motivation and students from high-quality schools may tend to "out perform" the model year after year. On the other hand, if the model includes a rich set of personal and community characteristics (represented by variable Z), the degree of correlation over time in the achievement errors may be minimal. In the analysis that follows, I maintain the assumption that errors e_1 and e_2 are uncorrelated. Some check on the validity of this assumption is provided by including in the model variables that measure individual motivation, attitudes, and postsecondary plans. In fact, I find that these variables are unrelated to achievement growth. I conclude from this finding that errors in achievement growth may reflect essentially random factors that differ from year to year. A more general investigation of this issue awaits the availability of data on at least three periods, rather than the two periods available in the High School and Beyond data.

The system of equations discussed above provides a convenient framework for assessing the bias to be expected from model misspecifications. Below, I will consider three issues: (1) the consequences, if any, of assessing the determinants of math proficiency using test data that span two or more years rather than a single year; (2) the consequences of replacing T_1 in equation (4.5) with proxies for prior achievement, such as personal characteristics and prior course enrollments; and (3) the consequences of excluding prior achievement (T_1) from equation (4.5).

The consequences of the first problem--missing an intermediate measure of performance--can be assessed by substituting the equation for T_1 (4.3) into the equation for T_2 (4.5), which yields:

$$(4.6) \quad T_2 = Z(\delta_2 + \theta_2\delta_1) + X_2\alpha_2 + X_1\alpha_1\theta_2 + \theta_1\theta_2T_0 + e_2 + \theta_2e_1$$

Equation (4.4) indicates that course enrollments X_2 are correlated with T_1 and therefore with e_1 from equation (4.3). As a result, the estimated coefficient on second-period enrollments (α_2) in (4.6) will be biased. In effect, within-period "shocks" to math proficiency (not accounted for by Z, X_1, and T_0), which have an opportunity to affect enrollments, will be mistakenly soaked up by period-2 enrollments. This effect is likely to be small in the HS&B data, which span only two years, since annual shocks to the level of mathematics proficiency are likely to be relatively small. This effect could, however, cause serious bias if the time between tests is long, say four or more years.[22] Equation (4.6) also reveals that the estimated influence of courses taken during the first half of a two-year period (X_1) may be exaggerated or diminished, depending on whether θ_2 is greater or less than one. Our empirical evidence suggests that θ_2 is quite close to unity. This factor is therefore not an important source of bias.

The consequences of replacing prior test score T_1 in equation (4.5) with the best available proxies can be assessed by substituting equation (4.1) into (4.6). This eliminates T_0 from the equation, yielding:

$$(4.7) \quad T_2 = Z(\delta_2 + \theta_2\delta_1 + \theta_1\theta_2\delta_0) + X_2\alpha_2 + X_1\alpha_1\theta_2 + e_2 + \theta_2e_1 + \theta_1\theta_2e_0.$$

As is evident, this equation, like the previous one, accords greater weight to personal characteristics to offset the absence of prior test information. Indeed, if the variance of e_0 and e_1 is zero--in other words, if personal characteristics (Z) perfectly predict math achievement scores T_0 and T_1--this equation will yield unbiased parameter estimates. Not surprisingly, this is an incredibly strong and unlikely assumption. If violated, the estimated effectiveness of period 1 and 2 course enrollments (coefficients α_1 and α_2) will be seriously biased due to their correlation with errors e_0 and e_1. The coefficients will reflect both the genuine effectiveness of these courses and the degree to which the course enrollments were influenced by prior achievement. Since many studies have been forced to rely on an equation similar to (4.7), I estimate a variant of it in Section V and compare it with results known to be unbiased under more general conditions. I will refer to estimator (4.7) as the prior achievement proxy model.[23]

Finally, consider the consequences of dropping period-1 enrollments from equation (4.7). Since enrollment decisions tend to be highly intercorrelated (e.g., calculus students typically enroll in a ninth through twelfth grade pattern of algebra 1, geometry, precalculus, and calculus), omitting X_1 from (4.7) clearly adds additional bias to estimates of the effect of personal characteristics (Z) and second-period enrollments on second-period learning. In Section V, we will refer to this model as the super biased model.

As a prelude to discussion of the consequences of measurement error in T_1, it is useful to write down alternative formulas for the parameters of the proxy and super biased models discussed above. First, express T_1 as functions of the right-hand-side variables included in each model:

(4.8) $T_1 + Z\pi_0 + X_1\pi_1 + X_2\pi_2 + \epsilon_\pi$

(4.9) $T_1 = Z\eta_0 + X_2\eta_2 + \epsilon_\eta$.

The coefficients π and η of these two "auxiliary" regressions have no direct structural interpretation but merely summarize the conditional relationships among the indicated variables. The effect of dropping the prior test control from equation (4.5) (the super biased model) or replacing it with proxies is found by substituting (4.9) and (4.8), respectively, into (4.5), yielding:

Prior Achievement Proxy Model

$$(4.10)\quad T_2 = Z(\delta_2 + \theta_2\pi_0) + X_1\pi_1\theta_2 + X_2(\alpha_2 + \theta_2\pi_2) + e_2 + \theta_2\epsilon_\pi$$

Super Biased Model

$$(4.11)\quad T_2 = Z(\delta_2 + \theta_2\eta_0) + X_2(\alpha_2 + \theta_2\eta_2) + e_2 + \theta_2\epsilon_\eta$$

In summary, this section presented a five-equation system of course enrollments and math learning. Two commonly used, but biased, estimators were derived and discussed within the context of a fully specified set of equations: the prior achievement proxy model and the super biased model. An important new finding also emerged: estimates of curricular learning equations may be subject to bias if outcomes (such as mathematics proficiency) are measured more than a year apart (assuming that enrollment decisions are made annually).[24] The HS&B data set fares reasonably well by this standard since achievement levels were measured only two years apart. The analysis suggests, however, that estimates of the effectiveness of courses taken primarily in twelfth grade may be slightly biased: positively biased for courses taken by upper-ability students (e.g., precalculus, calculus) and negatively biased for courses taken by lower-ability students. As previously mentioned, this type of bias increases as the time between outcome measurements increases and may become quite serious for outcomes measured four or more years apart. Pallas and Alexander's (1983) use of data with achievement levels measured at four-year intervals may account, in part, for their unexpected finding that nonadvanced math courses make no contribution to the development of mathematics skills.[25]

The Consequences of Measurement Error in Test Scores

As is well known, measurement error in an explanatory variable causes downward bias in the estimated regression coefficient associated with the error-ridden variable and either positive or negative bias in other parameters. Despite this, most studies discussed in the previous literature review failed to use estimators that dealt appropriately with measurement error. This may have been due to the fact that the studies were based on tests that were reasonably reliable (say, greater than 85 percent reliability). However, even modest levels of test error may generate severe bias if test scores are highly correlated with other variables included in the equation, such as, in the present case, course enrollments and personal characteristics, variables both highly correlated with achievement levels. Below, I develop two quite different strategies for solving the problem of measurement error in our period 1 test score (T_1).

As is customary, I assume that measured achievement T_1 is the sum of true, unobserved achievement (T_1^*) and an independent error (v):

$$(4.12) \qquad T_1 = T_1^* + \nu$$

Then, the measurement-error ratio associated with T_1 is given by:

$$(4.13) \qquad \lambda = \frac{\sigma_v^2}{\sigma_1^2}$$

where σ_v^2 = variance of the measurement error and σ_1^2 = variance of observed T_1. The reliability of T_1 is simply 1-λ. As previously mentioned, the consequences of measurement error in T_1 depend on the measurement-error ratio λ and the degree to which T_1 is correlated with the variables contained in equation (4.5), measured by the percentage variance explained in the auxiliary regression (equation 4.9 above) of T_1 on Z and X_2, that is, $R^2(T_1, Z, X_2)$.

The "effective" level of measurement error, then, is given by:

$$(4.14)\quad m = \frac{\lambda}{1 - R^2(T_1, Z, E_2)}$$

and the bias in OLS estimates of equation (4.5) caused by measurement error in T_1 is given by:

$$(4.15)\quad bias(\hat{\theta}_2) = E(\hat{\theta}_2 - \theta_2) = -\theta_2 m \leq 0$$

$$(4.16)\quad bias(\hat{\alpha}_2) = E(\hat{\alpha}_2 - \alpha_2) = \theta_2 \eta_2 m$$

$$(4.17)\quad bias(\hat{\delta}_2) = E(\hat{\delta}_2 - \delta_2) = \theta_2 \eta_0 m$$

where η_0 and η_2 are defined in auxiliary regression (4.9).[26] As is evident, the bias in OLS estimates of course effectiveness (α_2) is largest for courses taken by students with the highest prior achievement (represented by large values of η_2 from auxiliary regression (4.9)). As reported later in the paper in

Table 7, courses such as chemistry/physics, algebra 2, precalculus, and calculus have values of η_2 that range from 2.8 to 8.5 Even modest levels of measurement error are therefore likely to cause severe bias in the estimated effectiveness of these courses. If measurement error is quite large, the bias that it introduces is comparable to the bias introduced by totally excluding T_1 from equation (4.5). As the equations below indicate, OLS estimation of equation (4.5) yields parameter estimates that are weighted averages of the true structural parameters (θ_2, α_2, and δ_2) and the super biased coefficients given in equation (4.11):

$$(4.18)\quad E(\hat{\theta}_2) = (1-m)\theta_2 + m \cdot 0$$

$$(4.19)\quad E(\hat{\alpha}_2) = (1-m)\alpha_2 + m(\alpha_2 + \theta_2\eta_2)$$

$$=(1-m)\alpha_2 + m\alpha(superbiased)_2$$

$$(4.20)\quad E(\hat{\delta}_2) = (1-m)\delta_2 + m(\delta_2 + \theta_2\eta_0)$$

$$= (1-m)\delta_2 + m\delta(superbiased)_2 .$$

As the effective level of measurement m rises, the OLS estimates converge to the super biased estimates.

An important implication of this analysis is that including T_1 in the model of math learning without correcting for measurement error could result in estimated parameters that are no better, or even worse, than excluding T_1 and replacing it with reasonably good proxies! This is particularly likely to occur if the effective level of measurement error m is large (close to 1) and if enrollment choices X_1 and X_2 are highly partially correlated, given Z.[27] In this case, significant measurement error in T_1 permits X_2 to soak up the prior contribution of X_1 to T_1 and the effect of T_1 on X_2.

This analysis demonstrates the critical importance of controlling for test measurement error in curricular models of achievement growth. In more general models that allow mathematics learning to depend on math achievement T_1, math aptitude (A_1), and perhaps other achievement or aptitude measures, the problems caused by measurement error in several tests are similar to those analyzed above. The estimation techniques discussed in the next section extend readily to situations involving several tests measured with error.

A Consistent Estimator that Solves the Problem of Measurement Error

There are two major approaches to solving the problem of measurement error in explanatory variables such as prior math achievement. The first, which I will refer to as the errors in variables or EV estimator, uses estimates of the variances (and possibly covariances) of measurement errors to deflate (or attenuate) elements of the cross-product matrix of variables that are inflated due to measurement error. In the simple case, where only one explanatory variable is measured with error, for example, prior math achievement T_1 in model (4.5), this technique involves subtracting the estimated sum of squares of the measurement error υ_1 (that is, $E(\Sigma_i v_i^2) = N \cdot \sigma_v^2$, where N = the total sample size) from the sum of squares of observed prior math achievement T_{1i} (that is, $\Sigma_i T_{1i}^2$). This technique extends readily to situations with multiple variables measured with error.[28]

The EV estimator has not been widely used in the economics literature because estimates of the variances (and covariances) of measurement errors are typically unavailable in most data sets. In the case of test scores, however, these estimates are commonly available.

Arguably, the best, although least common, method for obtaining such estimates involves testing and retesting a group of students, ideally with the same or nearly identical test instruments. The two tests must be spaced far enough apart so that individual performance on the retest is not influenced by exposure to the first test. On the other hand, the tests must be spaced close enough together so that true achievement (or aptitude) is essentially identical for both tests. In other words, the tests must be measuring the same achievement (or aptitude) level.[29] Given two independent estimates of test performance, say test score T_i and retest score T_i', an estimate of the variance of measurement error in the test is given simply by $\sigma_v^2 = Var(T_i) - Cov(T_i, T_i')$.[30]

The more common method for estimating the variance of test measurement errors--and the one used to generate such information for the tests contained in the HS&B data set--is based on a single set of test scores. This technique generates an estimate of the reliability of a test referred to as Cronbach's coefficient alpha, or, in the special case of dichotomously scored test items (that is, right or wrong), Kuder-Richardson formula 20 (KR20) (Lord and Novick, 1968). This technique is comparable in spirit to the test-retest method since it measures the consistency of performance on different test items, as opposed to the consistency in performance across a test and its retest. As is well known, however, coefficient alpha is essentially a joint measure of the homogeneity of items on a test and individual consistency of performance across these items. Thus, to the extent that tests such as math proficiency tap more than one skill dimension, coefficient alpha will tend to understate a test's true reliability and therefore exaggerate its level of measurement error. On the other hand, to the extent that student test performance is influenced by day-specific factors such as alertness, mood, and testing conditions, coefficient alpha may actually overstate a test's true reliability. It would be fortuitous if the biases caused by test heterogeneity and day-specific error were generally offsetting, but I obviously have no way of knowing whether or not this is the case in the HS&B data.

Given the fact that coefficient alpha yields reliability estimates that may be overstated or understated, it follows that EV estimators constructed off such measures are subject to biases of unknown direction and size. In fact, Hoffer, Greeley, and Coleman (1985) and Jencks (1985) raise objections to Alexander and Pallas' (1985) use of this technique in the latter's analysis of the effectiveness of public and private schools. The former authors suggest that Pallas and Alexander (1985) used an exaggerated estimate of the severity of test measurement error (derived from coefficient alpha), thus causing downward bias in their EV estimate of the effectiveness of private schools.[31] As previously mentioned, however, the failure of coefficient alpha to reflect day-specific error suggests that the biases caused by this approach could be positive or negative. Given this uncertainty, it is fortunate that I am able to construct a second estimator that does not depend on an external estimate of the variance of the test error in measurement. This second estimator also offers an opportunity to assess in one particular data set the offsetting influences of test heterogeneity and day-specific error.

The second major approach to obtaining consistent parameter estimates in the presence of measurement error derives from the techniques developed to estimate simultaneous equation systems. In fact, the system defined by equation (4.1) and (4.5) can be used to construct a credible and powerful two-stage least squares (2SLS) estimator of mathematics learning. No additional assumptions are required other than those articulated above. The maintained assumption of zero correlation of the errors e_1 and e_2 is, of course, an assumption that would be desirable to relax for both the EV and 2SLS estimators.

The 2SLS estimator implied by equations (4.1) and (4.5) consists of a first-stage equation for prior achievement T_1, with right-hand-side instrumental variables X_1, X_2 and Z, and a second-stage equation consisting of T_1 predicted from the first stage, X_2 and Z. This technique essentially purges measured T_1 of its error component (v).

The reason this estimator "works" is that ninth and tenth grade enrollments do not directly affect learning in eleventh and twelfth grade and thus can be used to predict prior achievement measured during the spring of tenth grade.[32,33] As indicated in Section V, these instruments (i.e., X_1) add substantial explanatory power to the first-stage prior-achievement regression, above and beyond the variance explained by Z and X_2. As we shall see, this accounts for the excellent performance of the 2SLS estimator.

Despite the practical appeal of the 2SLS estimator discussed above, to my knowledge it has not previously been used in achievement-outcome studies. This may be explained in part by the fact that the estimator is identified because enrollment decisions during the first and second halves of high school (X_1 and X_2, respectively) are not perfectly linked. Some school-level inputs, on the other hand, may vary little over periods of two to four years, and thus may not provide sufficient variation to enable the 2SLS approach to work. The 2SLS approach could, however, be used to assess the effectiveness of high school classroom and teacher characteristics, which obviously vary from class to class. Such a model would need to take into account the fact that high school students are simultaneously exposed to different teachers and classrooms and different courses.[34] This suggests that any model of high school classroom and teacher effectiveness would need to embed within it a model of curricular effectiveness much like the one presented in this study. In fact, my curriculum model could be extended quite simply to allow for teacher and classroom effects by allowing the vector of course enrollment coefficients α_2 to vary as a function of classroom and teacher characteristics.

The previous analysis requires an important qualification. My conclusions are based on the maintained assumption of zero correlation between the errors in the achievement growth equations. If this assumption is false, both the EV and the 2SLS estimators will produce biased parameter estimates. In fact, if the achievement equation errors are correlated, prior course enrollments fail to qualify as valid instrumental variables for the 2SLS estimator. However, variants of the EV and 2SLS models discussed above can be designed to address both the problems of serial correlation and measurement error. As mentioned earlier, these models require data on at least three periods.

Summary

In this section I have examined the major econometric models used in studies of achievement outcomes. An important and perhaps surprising finding of the analysis is that curricular models of achievement growth that fail to correct for measurement error in prior achievement are likely to generate severely biased parameter estimates. These estimates may, in fact, be worse than the estimates obtained from models based only on proxies for prior achievement. I considered two estimators designed to solve the problem of measurement error in prior achievement, an errors in variables (EV) estimator and a two-stage least squares (2SLS) estimator.

V. PERFORMANCE OF ALTERNATIVE ESTIMATORS

This section compares the performance of alternative estimators of a curricular model of mathematics learning. My objective is to assess the extent to which common model misspecifications, of the kind discussed in Section IV, result in parameter estimates that are badly biased. Although the methodological findings presented in this section are limited, strictly speaking, to educational-outcome models that include course enrollments, they may also apply to educational-outcome models in general.

Since the focus of this section is on methodological issues, substantive interpretation of the empirical results is deferred to Section VI. Here, I present estimates for one particular population group: non-college-bound students. I focus on this group because there is strong substantive interest in the issue of improving their academic skills.

Summary of Alternative Estimators

Figure 1 provides a brief summary of the five major estimators discussed in Section IV. With the exception of the last estimator, the estimators are ranked in order of expected performance. The 2SLS and EV estimators provide the benchmark for comparing all other estimators. The 2SLS estimator, the EV estimator, and the estimator with no correction(s) for measurement error are estimated with and without controls for more than one achievement (or aptitude) score.

Two commonly used estimators, not discussed earlier, are also included in Figure 1: the simple least squares difference equation (#3) and an ad hoc estimator that I refer to as the residual growth specification (#7).[35] Both estimators impose particular values of θ_2 (the coefficient on prior math achievement T_1) in estimating structural equation (4.5). Any estimator that imposes θ_2 at some value, say $\tilde{\theta}$, shifts T_1 over to the left-hand side of the equation in the form of constructed variable $T_2 - \tilde{\theta}T_1$. This eliminates T_1 as a contaminating (bias-causing) source of measurement error in the equation, although measurement error adds to the error in the equation. This suggests that a simple least squares difference equation could generate estimates that are less biased than estimators based on missing or error-ridden measures of prior achievement (estimators 4 through 6 in Figure 1). Fortunately, there is no need to accept an estimator with any inconsistency if a data set contains information on prior achievement. If such data exist, the 2SLS or EV estimators dominate the simple difference equation, the residual growth specification, or any other estimator based on an arbitrarily imposed value of θ_2.[36]

Before comparing the alternative estimates presented in this section, it is useful to recognize that the seven estimators listed in Figure 1 can be viewed, more or less, as simple variations of a model in which a constructed dependent variable $T_2 - \tilde{\theta}T_1$ is regressed on Z and E_2. The

FIGURE 1

Summary of Alternative Estimators

Estimator		Dependent (Outcome) Variables	RHS Test Variables[a]	Other RHS Variables	Equation Number
1A	2SLS with single RHS test var.	ΔT or T_2	T_1	Z, X_2	--
1B	2SLS with multiple RHS test vars.	ΔT or T_2	T_1, A_1	Z, X_2	--
2A	EV with single RHS test var.	ΔT or T_2	T_1	Z, X_2	--
2B	EV with multiple RHS test vars.	ΔT or T_2	T_1, A_1	Z, X_2	--
3	Simple difference equation (OLS)	ΔT	None	Z, X_2	--
4A	Uncorrected measurement error with single RHS test var. (OLS)	ΔT or T_2	T_1	Z, X_2	(4.18) - (4.20)
4B	Uncorrected measurement error with multiple RHS test vars.	ΔT or T_2	T_1, A_1	Z, X_2	(4.18) - (4.20)
5	Prior achievement proxy model (OLS)	T_2	None	Z, X_1, X_2	(4.10)
6	Super biased model (OLS)	T_2	None	Z, X_2	(4.11)
7	Residual growth specification[b]	$T_2 - \hat{\rho}\, T_1$	None	Z, X_2	--

[a]If a model includes more than two tests, A_1 should be interpreted as a vector of test scores.

[b]This model is described in endnote 35.

consequences of alternative estimators (with different values of $\tilde{\theta}$) can then be found by subtracting equation (4.9) from (4.11), which yields:

$$(5.1)\quad T_2 - \tilde{\theta}T_1 = Z(\delta_2 + (\theta_2 - \tilde{\theta})\eta_0) + X_2(\alpha_2 + (\theta_2 - \tilde{\theta})\eta_2)$$

$$+ \epsilon_2 + (\theta_2 - \tilde{\theta})\epsilon_\eta .$$

Note that if $\tilde{\theta} = \theta_2$, the bias terms in (5.1) drop from the equation. The extreme model variants are represented by the super biased estimator ($\tilde{\theta} = 0$) and the simple difference equation ($\tilde{\theta} = 1$).[37] The implication of this analysis is that a model that generates badly biased estimates of θ_2 will also generate badly biased estimates of α_2 (course effectiveness) and δ_2 (the contribution of personal characteristics). In the empirical results presented below, I therefore focus on comparing a selected set of parameter estimates from alternative estimators, rather than all parameter estimates. These include $\hat{\theta}_2$ (the estimated coefficient on prior math achievement) and the coefficients on prealgebra, chemistry/physics, foreign languages, and calculus.

Is Measurement Error a Problem?

Table 6 presents estimates of the curricular model of mathematics learning for non-college-bound students using two different estimators, OLS with no correction for measurement error in T_1 (estimator 4A) and 2SLS (estimator 1A). First-stage estimates are reported in Table A-6.[38] Both models adopt a specification that includes only one right-hand-side test score. The biased least squares coefficient obtained on prior (sophomore) mathematics ($\hat{\theta}_2 = 0.68$) is similar to estimates obtained in previous studies. At face value, being so far from unity, this estimate indicates that mathematics skills depreciate quite rapidly over a two-year period. When compared with the 2SLS

TABLE 6
Estimates of the Simple Two-Stage Least Squares and Biased Least Squares Models of Math Gain for Non-College-Bound Students

Right-Hand-Side Variable	Biased OLS Model Coeff.	Biased OLS Model Std. Error	Simple 2SLS Model[a] Coeff.	Simple 2SLS Model[a] Std. Error
Sophomore math test	0.678	(0.011)	0.947	(0.029)
Credits in eleventh/twelfth grade				
Basic math	-1.555	(0.458)	-0.569	(0.495)
General math	-0.276	(0.249)	0.301	(0.270)
Computer math	-0.777	(0.803)	-1.705	(0.856)
Prealgebra	1.548	(0.588)	2.164	(0.626)
Algebra 1	2.370	(0.365)	2.503	(0.387)
Geometry	2.582	(0.303)	2.031	(0.326)
Algebra 2	2.891	(0.304)	1.893	(0.337)
Precalculus	2.357	(0.394)	1.512	(0.426)
Calculus	3.939	(1.239)	1.647	(1.332)
Specific vocational math	-0.205	(0.377)	0.026	(0.400)
Applied math	0.332	(0.287)	0.645	(0.305)
Chemistry/physics	1.583	(0.256)	0.828	(0.281)
Biology/survey science	-0.076	(0.176)	0.015	(0.186)
Math-related voc. ed.	0.427	(0.093)	0.205	(0.101)
Non-math-related voc. ed.	-0.107	(0.045)	-0.117	(0.048)
English/social studies	0.062	(0.065)	0.083	(0.069)
Foreign languages	0.434	(0.223)	0.112	(0.239)
Fine arts	-0.133	(0.093)	-0.146	(0.098)
Personal and other	-0.119	(0.087)	-0.143	(0.092)
Graduation indicator[b]	1.385	(0.307)	0.701	(0.332)
Female	-0.100	(0.164)	-0.901	(0.174)
Black	-1.586	(0.289)	-0.194	(0.355)
Hispanic	-1.569	(0.235)	-0.543	(0.268)
Asian	-1.245	(1.022)	-1.985	(1.088)
Suburban	0.183	(0.180)	0.008	(0.192)
Urban	0.028	(0.236)	-0.039	(0.251)
Northeast	-0.567	(0.234)	-0.584	(0.248)
West	0.453	(0.271)	0.353	(0.288)
South	-0.477	(0.205)	-0.117	(0.220)
Constant	2.860	(0.332)	0.934	(0.398)
R-square (T_2-T_1)	0.200		0.061	
Sample size	4,448		4,448	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.
[a]First-stage results for the sophomore math test equation are reported in Appendix Table A-6.
[b]The graduation indicator is set to 1 if the student graduated from high school, zero otherwise.

estimate of 0.95, however, it is clear that the biased least squares estimate is badly downwardly biased due to measurement error in the sophomore math test. The 2SLS estimate of θ_2 is actually quite close to unity. This suggests that math skills depreciate only modestly, if at all, for high school aged youth.[39]

I can compute an estimate of the level of measurement error in the sophomore mathematics test from these estimates, since equations (4.14) and (4.15) imply the following formulas for the effective degree of measurement error m and the measurement ratio λ:[40]

$$(5.2) \quad m = 1 - \hat{\theta}_2(OLS)/\hat{\theta}_2(2SLS)$$

$$(5.3) \quad \lambda = [1 - R^2(T_1, Z, X_2)] \cdot m$$

Given the R^2 reported in Table 7 below, this yields the following estimates:

$$\hat{m} = 1 - 0.716 = 0.284$$

$$\hat{\lambda} = [1 - 0.258] \cdot 0.284 = 0.211$$

As indicated, the effective level of measurement error (m) is about 1 1/3 times the measurement error ratio (λ). This increase is due to the significant collinearity between T_1 and the other variables in the math-learning equation (Z and X_2)--see Table 7.

The 2SLS estimate of λ is quite similar to the estimate obtained by Rock et al. (1985) using the psychometric technique discussed in Section IV (see Appendix Table A-1). As one would expect, both estimates are substantially larger than the estimated measurement-error ratio for <u>all</u> students, as reported in Appendix Table A-1 (1 - 0.868 = 0.132). This stems from the fact that the variance of sophomore test scores is much less in the restricted non-college-bound population than in the complete student population. One might expect, therefore, that the problem of measurement error would be exacerbated by estimating separate models for particular population groups (e.g., non-college-bound students). As explained below, however, this may or may not be the case since the level of

TABLE 7
Estimates of Auxiliary Sophomore
Math Test Regression for Non-College-Bound Students

Right-Hand-Side Variable	Coefficient	Standard Error
Credits in eleventh/twelfth grade		
Basic math	-3.666	(0.598)
General math	-2.147	(0.325)
Computer math	3.451	(1.051)
Prealgebra	-2.292	(0.770)
Algebra 1	-0.494	(0.478)
Geometry	2.049	(0.397)
Algebra 2	3.710	(0.395)
Precalculus	3.143	(0.515)
Calculus	8.521	(1.620)
Specific vocational math	-0.860	(0.493)
Applied math	-1.164	(0.375)
Chemistry/physics	2.808	(0.332)
Biology/survey science	-0.336	(0.230)
Math-related voc. ed.	0.827	(0.121)
Non-math-related voc. ed.	0.038	(0.059)
English/social studies	-0.075	(0.085)
Foreign languages	1.200	(0.292)
Fine arts	0.048	(0.121)
Personal and other	0.089	(0.113)
Graduation indicator	2.544	(0.400)
Female	-0.361	(0.215)
Black	-5.177	(0.371)
Asian	-3.892	(1.339)
Hispanic	-3.818	(0.302)
Urban	0.251	(0.310)
Suburban	0.651	(0.236)
Northeast	0.064	(0.306)
West	0.373	(0.356)
South	-1.339	(0.268)
Constant	7.165	(0.422)
R-square	0.258	
Sample size	4,448	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.

collinearity between T_1, Z, and X_2 tends to be significantly lower in restricted samples, thereby offsetting the increase in relative measurement error. This point is illustrated below.

Define $\hat{T}_1$ as the component of T_1 explained by Z and E_2, and $\in$ as the unexplained part of T_1. As discussed in Section IV, the effective degree of measurement error depends on the measurement-error ratio $\lambda = \sigma_v^2/Var(T_1)$ and the collinearity between T_1, Z, and X_2, as measured by $R^2(T_1, Z, X_2) = 1 - Var(\in)/Var(T_1)$. Then, the effective degree of measurement error, given a selection rule S, is given by

$$(m|S) = \frac{(\lambda|S)}{1 - R^2(T_1, Z, X_2|S)}$$

$$= \frac{\sigma_v^2/Var(T_1|S)}{Var(\epsilon|S)/Var(T_1|S)}$$

$$= \frac{\sigma_v^2}{Var(\epsilon|S)}.$$

As is evident, exogenous selection rules that are uncorrelated with $\in$ leave the effective degree of measurement error unchanged. On the other hand, endogenous selection rules may alter the variance of ($\in$/S), thereby changing the effective degree of measurement error. Since my selection rules (stratification by predicted prior math achievement and college-bound and non-college-bound) are based on exogenous variables X_1, X_2, and Z, the problem of measurement error may be made neither better nor worse by splitting the data into subgroups.

As indicated in Table 6, measurement error in T_1 has a profound effect on the biased least squares estimates. In particular, the least squares estimate of the calculus coefficient is more than double the 2SLS estimate. The chemistry/physics, precalculus, and algebra 2 coefficients are similarly upwardly biased. In contrast, the coefficients on prealgebra and non-math-related vocational education are downwardly biased. In general, measurement error in T_1 generates strong positive bias for those courses taken by students with the highest prior math scores and negative bias for courses taken by students with the lowest prior math scores. The coefficients on courses taken in eleventh or twelfth grade by mid-level students--algebra 1 and English, for example--are essentially unaffected by measurement error. These results accord, as they should, with the estimates in Table 7 of the auxiliary regression of T_1 on X_2 and Z. (See equation 4.19.)

One additional indication of misspecification in the biased least squares estimator is the fact that the foreign language coefficient is rather substantial (0.434), particularly when compared with the 2SLS estimate (0.112 with a standard error of 0.239). A priori, it is difficult to believe that foreign language instruction contributes significantly to mathematics learning. Indeed, the 2SLS estimate indicates that it does not.

At face value, the 2SLS estimates suggest that the most productive high school math course is algebra 1, followed (in order) by prealgebra, geometry, and algebra 2. As discussed in Section III, this pattern of results may be explained, at least in part, by the fact that the HS&B test includes items that require algebra and geometry skills. If so, it could be argued that the HS&B test is not the perfect test for measuring the contribution to mathematics learning of courses such as algebra 2, precalculus, chemistry, and math-related vocational education. Since eleventh and twelfth grade enrollments in these courses substantially outweigh enrollments in prealgebra, algebra 1, and geometry (see Table 3), it might have been better if the HS&B math test had placed greater emphasis on measuring cognitive mathematics skills and/or specific achievement skills related to the subjects taken during the eleventh and twelfth grade. As a result, it is important to interpret my empirical estimates with some caution. Clearly, it would be useful to reestimate the models presented in this study using a richer and more extensive battery of mathematics tests.[41]

The overall performance of the 2SLS estimator is excellent. The coefficient on prior mathematics achievement is very precisely estimated and, as a result, the estimated standard errors in the 2SLS model are only slightly larger than the OLS estimates. This fortunate result is due to the fact that the unique exogenous variables in the first-stage equation (i.e., those variables that are excluded from the second-stage equation) add substantial "kick" to this equation--the R^2 increases from 26 percent in Table 7 to 39 percent in Appendix Table A-6.

Despite the excellent performance of the model, it explains a relatively small share of the variance in achievement growth, as measured by the $R^2(T_2 - T_1)$ statistic. Note that the tables in this paper report the R^2 statistic for the model with T_2-T_1 as the dependent variable rather than T_2. The R^2 statistic for the model with T_2 is much higher. It also provides a misleading sense of the predictive power of the model because virtually all of the explanatory power in that model is accounted for by prior achievement. However, to maintain consistency with the text, the coefficient on prior math achievement is taken from the model with T_2 as the dependent variable. The comparable coefficient for a model with a dependent variable of (T_2-T_1) is given by the reported coefficient minus one. In other respects, the two forms of the model are identical. The low explanatory power of the model is due in part to the fact that the reliability of (T_2-T_1) (i.e., test gain) is only 30 percent (Appendix Table A-1). In fact, the 2SLS model explains 20 percent of the "usable" variance of (T_2-T_1).[42] This still indicates, however, that idiosyncratic factors are a major determinant of individual achievement growth, at least as measured by the HS&B test for the period covering eleventh and twelfth grades. The explanatory power of the model might have been higher for a test more closely matched to the mathematics content of eleventh and twelfth grades, for example, algebra 2, precalculus, calculus, chemistry, and physics.

In order to provide some check on the validity of the maintained assumption that the achievement equation errors are uncorrelated over time, I reestimated the model with the following additional variables: a measure of individuals' expected schooling at the end of tenth grade, two attitudinal variables that measure whether "students have math anxiety or a fear of math" and whether "students believe that math is important or useful," and a measure of mother's educational attainment, as reported by students.[43] The estimates indicate, perhaps surprisingly, that these variables are unrelated to achievement growth, conditional on prior achievement, course enrollments, and basic demographic characteristics. I conclude from this finding that unexplained achievement growth may be due to essentially random factors that differ from year to year. As mentioned earlier, a more general investigation awaits the availability of data containing at least three periods.

Alternative 2SLS and EV Estimates

Table 8 presents estimates of three alternative estimators of the model of mathematics learning: (1) a general 2SLS estimator that includes prior science and verbal test scores, as well as prior mathematics achievement (estimator 1B); (2) a general EV estimator comparable to this estimator (estimator 2B); and (3) a simple EV estimator comparable to the simple 2SLS estimator discussed above (estimator 2A). The respective estimators "correct" for measurement error in all right-hand-side test scores.

The reported standard errors for the 2SLS model are the conventional asymptotic standard errors for the 2SLS estimator. The standard errors for the EV models are based on the regression output of a standard ordinary least squares package. As mentioned earlier, the EV estimates were derived from an adjusted cross-product matrix. The matrix was adjusted to correct for the presence of measurement error in the variance of all achievement scores. As pointed out by Fuller (1987), the standard errors obtained from an ordinary least squares package are biased downward for EV models. Fuller (1987) presents correct asymptotic formulas for the variance of EV parameter estimates. Since

TABLE 8
Estimates of Alternative Two-Stage Least Squares and Errors in Variables Models of Math Gain for Non-College-Bound Students

Right-Hand-Side Variable	General 2SLS Model with 3 Endogenous Soph. Tests [a] Coeff.	Std. Error	General EV Model with 3 Endogenous Soph. Tests Coeff.	Std. Error	EV Model with Endogenous Soph. Math Test Coeff.	Std. Error
Sophomore math test[b]	1.057	(0.071)	0.946	(0.020)	0.913	(0.014)
Sophomore science test[b]	-0.159	(0.219)	-0.098	(0.058)		
Sophomore verbal test[b]	0.076	(0.095)	-0.003	(0.025)		
Credits in eleventh/twelfth grade						
Basic math	-0.620	(0.504)	-0.726	(0.501)	-0.710	(0.501)
General math	0.216	(0.286)	0.123	(0.274)	0.158	(0.274)
Computer math	-1.594	(0.870)	-1.595	(0.869)	-1.595	(0.869)
Prealgebra	2.199	(0.637)	2.106	(0.639)	2.055	(0.638)
Algebra 1	2.461	(0.402)	2.433	(0.400)	2.389	(0.400)
Geometry	2.034	(0.330)	2.086	(0.330)	2.080	(0.330)
Algebra 2	1.888	(0.347)	2.092	(0.332)	2.063	(0.331)
Precalculus	1.538	(0.439)	1.577	(0.428)	1.627	(0.428)
Calculus	1.635	(1.338)	1.940	(1.342)	1.939	(1.343)
Specific voc. math	-0.001	(0.405)	0.024	(0.408)	0.007	(0.408)
Applied math	0.575	(0.310)	0.566	(0.315)	0.556	(0.315)
Chemistry/physics	0.884	(0.296)	0.963	(0.278)	0.926	(0.278)
Biology/survey science	0.031	(0.195)	0.027	(0.192)	-0.000	(0.191)
Math-related voc. ed.	0.257	(0.115)	0.266	(0.101)	0.245	(0.101)
Non-math-related voc. ed.	-0.115	(0.048)	-0.113	(0.049)	-0.114	(0.049)
English/social studies	0.084	(0.071)	0.094	(0.071)	0.089	(0.070)
Foreign languages	0.056	(0.250)	0.152	(0.243)	0.154	(0.242)
Fine arts	-0.134	(0.101)	-0.130	(0.101)	-0.141	(0.101)
Personal and other	-0.159	(0.095)	-0.155	(0.094)	-0.144	(0.094)
Graduation indicator	0.694	(0.335)	0.760	(0.335)	0.763	(0.335)
Female	-1.038	(0.291)	-1.041	(0.187)	-0.926	(0.178)
Black	-0.339	(0.492)	-0.508	(0.327)	-0.338	(0.319)
Hispanic	-0.560	(0.396)	-0.771	(0.263)	-0.633	(0.257)
Asian	-0.292	(1.163)	-0.537	(1.109)	-0.341	(1.107)
Suburban	-0.055	(0.199)	-0.026	(0.196)	-0.002	(0.196)
Urban	-0.123	(0.280)	-0.051	(0.259)	-0.005	(0.257)
Northeast	-0.592	(0.250)	-0.612	(0.254)	-0.598	(0.254)
West	0.323	(0.308)	0.423	(0.296)	0.404	(0.295)
South	-0.158	(0.227)	-0.220	(0.223)	-0.198	(0.223)
Constant	1.420	(1.204)	1.7606	(0.443)	1.189	(0.364)
R-square (T_2-T_1)	0.060		0.067		0.065	
Sample size	4,389		4,410		4,410	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.
[a]First-stage results are reported in Appendix Table A-6.
[b]Endogenous variables subject to measurement error.

the EV estimates in this paper are presented for comparative purposes only, I have not taken the extra step of computing asymptotic standard errors for the estimates.

As is evident in Table 8, the three alternative estimators yield results that are strikingly similar to the simple 2SLS estimates in Table 6 discussed above. For example, the chemistry/physics coefficients are, respectively, 0.83, 0.88, 0.96, and 0.93 in these four models. The precalculus coefficients are, respectively, 1.51, 1.54, 1.58, and 1.63.[44] There is no strong evidence that mathematics learning is affected by prior skills other than mathematics. As one would expect, however, the standard errors on the sophomore math, science, and verbal coefficients in the general 2SLS model are much larger than in the simple 2SLS model. This is due to the fact that the same instrumental variables are used to predict all three sophomore test scores in their first-stage equations.

Earlier in this section, I observed that the psychometric estimate of the level of measurement error in T_1 (λ) was nearly identical to the estimate derived from the 2SLS estimator. As a result, we should not be too surprised that our 2SLS and EV estimators generate nearly identical estimates. In other contexts, of course, the two estimators may yield quite different results. On the basis of this study, however, one could argue that it would generally be better to use an EV estimator to correct for measurement error in test scores than to ignore measurement error altogether.

How Bad Are the Alternative Least Squares Estimators?

This section assesses the performance of the six commonly used estimators summarized in Figure 1. Estimates of selected parameters are presented in Table 9 for each estimator.[45] The estimators are ordered, for the most part, in terms of their respective estimates of θ_2, the coefficient on prior math achievement. Hence, results for the simple difference equation appear in the first row and results for the super biased model appear in the last row.

The principal conclusion to be drawn from Table 9 is that the estimators that have dominated previous studies of educational outcomes--the least squares model with no measurement-error

TABLE 9
The Effect of Alternative Estimators on Selected Parameter Estimates
(for Non-College-Bound Students)

Estimator		Sophomore Math Test[a]	Prealgebra	Chemistry/Physics	Foreign Languages	Calculus
3	Simple difference equation	1.000[a]	2.285	0.680	0.048	1.197
1A	2SLS with single RHS test	0.947 (0.029)	2.164 (0.626)	0.828 (0.281)	0.112 (0.239)	1.647 (1.332)
1B	2SLS with multiple RHS tests	1.057	2.199	0.884	0.056	1.635
2A	EV with single RHS test	0.913	2.055	0.926	0.154	1.939
2B	EV with multiple RHS tests	0.946	2.106	0.963	0.152	1.940
4A	Uncorrected measurement error in single tests	0.678	1.548	1.583	0.434	3.939
4B	Uncorrected measurement error in multiple tests	0.600	1.500	1.434	0.340	3.623
5	Prior achievement model	0.430[c]	0.897	2.465	0.315	6.039
6	Super biased model	0.000[b]	(0.007)	3.488	1.248	9.718

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.

Notes: Estimated standard errors are reported only for estimator 1A, since the estimates vary only slightly across the different estimators. The estimates reported in this table were drawn from the complete model estimates reported in Tables 6, A-7, and A-8.

[a] The sophomore math test coefficient is fixed at a value of 1.0 in the simple difference equation estimator.
[b] The sophomore math test coefficient is fixed at a value of 0.0 in the super biased model.
[c] This estimate is intended to illustrate the implicit level of bias caused by replacing T_1 with proxy variables. The procedure used to construct this estimate is discussed in endnote 46.

correction (#4), the prior achievement proxy model (#5), and the super biased model (#6)--perform atrociously within the context of my curricular model of mathematics learning. With respect to estimates of θ_2, these estimators deviate radically from the simple 2SLS estimate of 0.95. The prior achievement proxy and super biased models are particularly bad: the implied value of θ_2 in the proxy model is 0.430 and the imposed value of $\hat{\theta}_2$ in the super biased model is, of course, zero.[46] As expected, these estimators also yield estimates of curricular effectiveness that are severely distorted. The calculus coefficient, in particular, is terrifically sensitive to incomplete control for prior achievement, as is the coefficient on foreign language coursework. In the super biased model, for example, the calculus coefficient is inflated by a factor of 590 percent and the foreign language coefficient is inflated by a factor of 1,114 percent over the simple 2SLS coefficients (Table 6). In the prior achievement proxy model, the comparable figures are 239 percent and 388 percent, respectively.

For future studies that may lack the requisite data to use one of the EV or 2SLS estimators, it may be helpful to note that if we believe that foreign language instruction makes no contribution to mathematics learning, a hypothesis that is confirmed by our 2SLS estimates, the foreign language coefficient actually is a very sensitive measure of model misspecification. In the present context, the alternative least squares estimates (aside from the simple difference equation) are emphatically rejected by this "specification test."

<u>Summary</u>

In this section I demonstrated the importance of adequately controlling for prior mathematics achievement when estimating the value-added contribution of high school coursework to mathematics proficiency. Moreover, I found that apparently slight imperfections in measured prior achievement dramatically affected estimates of curricular effectiveness. I successfully demonstrated the merits of the structural 2SLS estimator that was discussed in Section IV and found that alternative EV estimates were remarkably similar to 2SLS estimates. It would be interesting in future studies to see if EV estimators yield statistical results that are close to estimates that do not depend on external estimates of the level of measurement error.

My results suggest that empirical studies of curricular effectiveness must be based on longitudinal outcome data. The estimators that were based only on senior mathematics achievement were found to generate quite unsatisfactory results. This means that large and expensive data bases such as the National Assessment of Educational Progress are not useful for exploring the determinants of achievement growth. Since explorations of this type are of vital importance, it seems obvious that greater national attention should be given to developing longitudinal data bases that, like the High School and Beyond study, include extensive achievement, student, and school-level data.

The next section evaluates and interprets the empirical estimates obtained using the 2SLS estimator. The estimates from the EV estimator were quite similar and therefore are not discussed.

VI. MODEL ESTIMATES FOR ACADEMICALLY DISADVANTAGED AND ADVANTAGED STUDENTS

The objectives of this section are to: (1) assess the extent to which mathematics learning is substantial in courses other than traditional mathematics courses, in particular, math-related vocational education, applied/vocational math, and math-related science; (2) estimate the extent to which the effectiveness of mathematics and math-related courses varies among different student groups; and (3) assess the principle that mathematics can be learned in an applied context--for example, in a science lab, or in a vocational workshop.

In order to explore the possible differential effectiveness of courses, I present separate estimates of the 2SLS model for college-bound and non-college-bound students and for students classified into sophomore math proficiency triptiles (thirds). These two variables were used to stratify the student population because previous studies have shown that they are the two most important predictors of math enrollment decisions. Students are classified into math proficiency triptiles on the basis of their predicted rather than actual sophomore math score in order to avoid inducing truncation (or selection) bias into the estimates for each triptile.[47] This is necessary because predicted rather than actual prior achievement appears in the second stage of the 2SLS estimator. If prior achievement were measured without error, incidentally, it would be perfectly legitimate to stratify the sample on the basis of actual prior achievement, as long as this variable is included as an explanatory variable. As indicated in Appendix Table A-5, the sophomore math regression used to predict sophomore math scores explains a large share (55 percent) of the variable's "usable" variance--its reliability is 87 percent, as indicated in Appendix Table A-1. If otherwise, it would be a poor variable for separating the population into distinctly different groups.

Separate estimates for college-bound students and all students are presented in Table 10. (The comparable estimates for non-college-bound students were given in Table 6.) Estimates by math triptiles are given in Table 11. These estimates are based on the simple 2SLS estimator (estimator 1A). Virtually identical results were obtained from using a general 2SLS estimator that includes prior science and verbal achievement as well as prior math achievement (estimator 1B). These alternative estimates are contained in Appendix Tables A-9 and A-10.

As indicated in Tables 10 and 11, the evidence that mathematics skills can and are being learned outside of traditional mathematics courses is striking. Although the specific estimated effects vary across population groups (in part, due to finite sample variation), specific vocational math, applied math, and chemistry/physics all contribute substantially to the development of mathematics proficiency, particularly for the top two triptiles. In fact, for the middle triptile, math-related science is nearly 50 percent as productive as typical intermediate to advanced math courses in promoting

TABLE 10
Estimates of Simple 2SLS Model of Math Gain for College-Bound Students and All Students[a]

Right-Hand-Side Variable	College-Bound Students Coeff.	College-Bound Students Std. Error	All Students Coeff.	All Students Std. Error
Sophomore math test[b]	1.024	(0.026)	0.997	(0.018)
Credits in eleventh/twelfth grade				
Basic math	1.644	(0.780)	0.044	(0.406)
General math	1.044	(0.276)	0.743	(0.190)
Computer math	-1.293	(0.423)	-1.291	(0.386)
Prealgebra	1.507	(0.431)	1.790	(0.358)
Algebra 1	2.271	(0.277)	2.406	(0.225)
Geometry	1.236	(0.186)	1.427	(0.161)
Algebra 2	1.671	(0.156)	1.728	(0.142)
Precalculus	0.969	(0.188)	1.115	(0.169)
Calculus	0.441	(0.283)	0.726	(0.269)
Specific vocational math	1.605	(0.441)	0.679	(0.293)
Applied math	1.403	(0.378)	0.953	(0.233)
Chemistry/physics	0.473	(0.123)	0.613	(0.112)
Biology/survey science	-0.164	(0.133)	-0.084	(0.108)
Math-related voc. ed.	-0.085	(0.099)	0.069	(0.070)
Non-math-related voc. ed.	-0.145	(0.058)	-0.136	(0.036)
English/social studies	-0.144	(0.062)	-0.029	(0.045)
Foreign languages	0.063	(0.102)	0.153	(0.094)
Fine arts	-0.089	(0.071)	-0.075	(0.057)
Personal and other	-0.151	(0.073)	-0.142	(0.057)
Graduation indicator	0.669	(0.503)	0.757	(0.265)
Female	-0.420	(0.148)	-0.639	(0.111)
Black	-0.155	(0.293)	-0.161	(0.216)
Asian	-0.845	(0.561)	-0.716	(0.508)
Hispanic	-0.351	(0.294)	-0.381	(0.196)
Urban	0.266	(0.212)	0.097	(0.161)
Suburban	0.151	(0.164)	0.081	(0.124)
Northeast	0.611	(0.207)	0.079	(0.158)
West	0.118	(0.225)	0.219	(0.177)
South	-0.262	(0.191)	-0.172	(0.144)
Constant	0.809	(0.626)	0.634	(0.310)
R-square (T_2-T_1)	0.080		0.085	
Sample size	6,472		10,960	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.
[a]First-stage results for each group are reported in an appendix available from the author. The variables included in each first stage are identical to those listed in Appendix Table A-6.
[b]Endogenous variables subject to measurement error.

TABLE 11
Estimates of Simple 2SLS Model of Math Gain for Each Predicted Sophomore Math Test Triptile[a]

Right-Hand-Side Variable	Lower Third		Middle Third		Upper Third	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Sophomore math test[b]	0.991	(0.071)	0.909	(0.047)	0.913	(0.035)
Credits in eleventh/twelfth grade						
Basic math	-0.453	(0.473)	3.125	(1.307)	2.985	(2.302)
General math	0.436	(0.288)	0.782	(0.346)	1.429	(0.408)
Computer math	-2.626	(2.251)	-1.461	(0.902)	-1.163	(0.380)
Prealgebra	1.749	(0.584)	1.570	(0.570)	2.342	(0.734)
Algebra 1	2.276	(0.358)	2.617	(0.391)	2.014	(0.425)
Geometry	1.406	(0.509)	1.027	(0.238)	1.839	(0.252)
Algebra 2	1.992	(0.548)	2.080	(0.265)	1.592	(0.170)
Precalculus	-0.427	(1.005)	1.797	(0.445)	1.284	(0.177)
Calculus	1.865	(7.036)	3.974	(1.869)	1.055	(0.271)
Specific voc. math	-0.045	(0.417)	1.110	(0.495)	2.379	(0.781)
Applied math	0.632	(0.305)	1.187	(0.442)	1.816	(0.712)
Chemistry/physics	0.165	(0.488)	0.944	(0.233)	0.576	(0.122)
Biology/survey science	0.111	(0.200)	-0.179	(0.200)	-0.144	(0.157)
Math-related voc. ed.	0.369	(0.146)	0.016	(0.111)	-0.136	(0.107)
Non-math-related voc. ed.	-0.158	(0.053)	-0.179	(0.061)	0.042	(0.086)
English/social studies	0.057	(0.076)	-0.079	(0.080)	-0.108	(0.084)
Foreign languages	0.144	(0.307)	-0.037	(0.179)	0.239	(0.108)
Fine arts	-0.059	(0.123)	-0.126	(0.096)	-0.038	(0.082)
Personal and other	-0.147	(0.106)	-0.165	(0.099)	-0.177	(0.087)
Graduation indicator	0.590	(0.336)	0.852	(0.592)	1.850	(1.345)
Female	-0.756	(0.205)	-0.764	(0.198)	-0.501	(0.174)
Black	-0.324	(0.324)	-0.273	(0.449)	0.243	(0.615)
Asian	-1.640	(1.399)	-0.910	(0.906)	-0.439	(0.604)
Hispanic	-0.325	(0.278)	-1.045	(0.389)	0.233	(0.485)
Urban	0.168	(0.267)	0.074	(0.288)	0.101	(0.268)
Suburban	0.263	(0.228)	-0.270	(0.218)	0.363	(0.190)
Northeast	0.165	(0.319)	-0.117	(0.284)	0.033	(0.218)
West	0.711	(0.360)	-0.041	(0.292)	0.218	(0.261)
South	0.113	(0.256)	-0.256	(0.249)	-0.505	(0.231)
Constant	0.330	(0.518)	2.387	(0.812)	1.345	(1.608)
R-square (T_2-T_1)	0.044		0.073		0.076	
Sample size	3,647		3,646		3,665	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.
[a]First-stage results for each triptile are reported in an appendix available from the author. The variables included in each first stage are identical to those listed in Appendix Table A-6. Triptile samples are described in the text.
[b]Endogenous variables subject to measurement error.

math-skills development. The applied math courses are even more productive. For the top two triptiles, specific vocational and applied math courses contribute from 1.1 to 2.4 points to achievement growth.

The evidence for the lower triptile provides modest support for the notion that applied courses contribute to math achievement, although the specific pattern of coefficients differs somewhat. In particular, math-related vocational education makes an important contribution to math learning for this group. As one would expect, the effect of a single math-related vocational course (0.369) is less than the effect of a traditional math course; for example, the prealgebra coefficient is 1.749 for this group. A more informative comparison, however, should take into account the fact that students in this group take substantial amounts of vocational education, in fact, 7.7 times as much vocational education as mathematics in grades eleven and twelve (see Table A-4). Thus, the cumulative effect of enrollments in math-related vocational courses could be quite substantial. If, for example, students take two credits of math-related vocational education in eleventh grade and three credits of math-related vocational education in twelfth grade--a typical enrollment pattern for a vocational concentrator--the gain in math achievement from these courses alone would amount to 1.35 points.[48] In fact, the estimates suggest that a portfolio of five credits of math-related vocational education and two applied math credits would raise math achievement for the lower third by 3.2 points, almost four times their actual gain of 0.85 points.[49] Additional enrollments in math-related science or traditional mathematics could, of course, add substantially to this total. Indeed, Table 11 indicates that prealgebra, algebra 1, geometry, and algebra 2 are just as effective for students in the lower triptile as for those students in the top two triptiles. However, the effectiveness of math-related vocational education is limited to non-college-bound students (Table 6) and students in the lower triptile. This implies that the mathematical content of these courses is, at present, rather elementary.

At face value, the estimates indicate that specific vocational math provides no contribution to math proficiency for non-college-bound students and students in the lowest triptile. The effectiveness of applied math for these students is stronger (about 0.65), but still less than the estimates for college-bound students (Table 10). These estimates suggest that non-college-bound students may be enrolled in different types of specific vocational and applied math courses than college-bound students. In particular, the applied math courses taken by non-college-bound students may be less mathematically challenging than the applied math courses taken by college-bound students.

One puzzling aspect of Tables 10 and 11 is the fact that the estimated effectiveness of upper-level courses--calculus, chemistry/physics, precalculus, and algebra 2--is lower for the top triptile than for the middle triptile and lower for college-bound students than for non-college-bound students. This result could be an artifact due to the fact that the top triptile is more likely to have students who received a perfect or near-perfect test score at the end of grade 12. One way to examine this possibility is to toss out the top 10 percent of the sample, based on the predicted sophomore math score. Table 12 presents estimates for the group in the fiftieth to ninetieth percentiles. With the top 10 percent excluded, I find, as surmised, that the coefficients for the advanced courses listed above jumped to levels quite comparable to estimates obtained for the middle third.

VII. CONCLUSION

My analysis demonstrates, as expected, that participation in mathematics courses significantly enhances mathematics proficiency. This finding lends support to the common sense recommendation in A Nation at Risk that high school students should be required to take additional mathematics courses. But, the analysis also demonstrates that the development of mathematics skills is substantial in certain kinds of vocational-technical courses, quantitatively oriented science courses such as

TABLE 12
Estimates of Simple 2SLS Model of Math Gain for Students with Predicted Sophomore Math Scores between the 50th and 90th Percentiles[a]

Right-Hand-Side Variable	Coefficient	Standard Error
Sophomore math test[b]	0.976	(0.036)
Credits in eleventh/twelfth grade		
Basic math	3.427	(1.371)
General math	1.589	(0.386)
Computer math	-1.405	(0.552)
Prealgebra	3.171	(0.629)
Algebra 1	2.487	(0.374)
Geometry	1.482	(0.211)
Algebra 2	2.008	(0.172)
Precalculus	1.579	(0.216)
Calculus	1.601	(0.558)
Specific vocational math	0.962	(0.569)
Applied math	1.219	(0.547)
Chemistry/physics	0.930	(0.140)
Biology/survey science	-0.223	(0.161)
Math-related voc. ed.	0.066	(0.100)
Non-math-related voc. ed.	-0.056	(0.064)
English/social studies	-0.051	(0.076)
Foreign languages	0.108	(0.122)
Fine arts	-0.066	(0.079)
Personal and other	-0.254	(0.087)
Graduation indicator	0.362	(0.810)
Female	-0.723	(0.170)
Black	-0.249	(0.470)
Asian	-0.589	(0.679)
Hispanic	-0.551	(0.421)
Urban	0.418	(0.258)
Suburban	0.058	(0.187)
Northeast	0.204	(0.233)
West	0.414	(0.255)
South	-0.323	(0.221)
Constant	1.062	(1.012)
R-square (T_2-T_1)	0.111	
Sample size	4,381	

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.
[a]First-stage results for each group are reported in an appendix available from the author. The variables included in each first stage are identical to those listed in Appendix Table A-6.
[b]Endogenous variables subject to measurement error.

chemistry and physics, and applied math courses such as business and consumer mathematics. In other words, learning mathematics in an applied context is a viable alternative or complement to enrolling in traditional mathematics. This finding is particularly important for non-college-bound and academically disadvantaged students, who tend to take the minimum amount of required mathematics but substantial amounts of vocational education. It is also important for college-bound students, a group that actually takes more vocational education than mathematics in high school. The estimates suggest, however, the need to improve the mathematical rigor of the math-related vocational courses and specific vocational math courses taken by non-college-bound students.

This suggests that the next wave of academic reforms should encourage systemic reform of much of the high school curriculum, in particular, vocational education, science, and mathematics. The current reforms implicitly and mistakenly assume that mathematics instruction is the sole province and responsibility of high school mathematics departments, whereas the empirical results reported in this paper demonstrate that vocational education and science could be important vehicles for teaching mathematics. Although the empirical findings of this study pertain only to the development of mathematics skills, it is plausible that the more general principle is also true, namely that academic skills such as communications, problem solving, and mathematics can be learned in an applied context, outside of traditional English and mathematics courses. If so, future academic reforms should consider the extent to which subjects such as vocational education, science, fine arts, and social studies can extend, reinforce, and motivate interest in academic skills that have traditionally been acquired primarily in English and traditional math courses. Such reforms will undoubtedly prove more difficult to design and implement than simply changing high school graduation requirements. However, given the fact that minimum graduation requirements in mathematics and science have not risen above two courses (in each area) in the vast majority of states, it would be prudent to broaden the arsenal of policy tools used to stimulate growth in academic competencies among high school students.

TABLE A-1

The Mean, Variance, and Estimated Reliability of the HS&B Sophomore Cohort Test Battery (Formula Scores)

Test	Number of Test Items	Error Variance[a]	All Students Mean	All Students Variance	All Students Reliability	Non-College-Bound Students Mean	Non-College-Bound Students Variance	Non-College-Bound Students Reliability
Sophomore test								
Mathematics	38	12.390	13.436	94.090	0.868	8.937	64.697	0.808
Verbal[b]	40	10.768	16.146	86.751	0.876	11.834	60.058	0.821
Reading	19	5.198	7.167	22.677	0.771	5.118	15.999	0.675
Vocabulary	21	5.570	8.973	28.016	0.801	6.713	21.345	0.739
Science	20	5.570	9.255	19.945	0.721	7.663	17.646	0.684
Writing[c]	17	5.290	8.810	24.701	0.786	6.711	22.790	0.768
Civics[c]	10	3.423	4.715	6.970	0.509	3.789	6.252	0.453
Senior test								
Mathematics		11.834	15.354	113.401	0.896	9.908	73.945	0.840
Verbal[b]		10.170	19.568	99.421	0.898	11.834	60.058	0.831
Reading		5.062	8.411	25.928	0.805	6.161	18.623	0.728
Vocabulary		5.108	11.154	31.047	0.835	8.694	25.652	0.801
Science		5.336	10.201	20.648	0.742	8.181	17.765	0.700
Writing[c]		4.709	10.491	24.265	0.806	8.387	24.772	0.810
Civics[c]		3.133	5.828	7.506	0.583	4.790	7.186	0.564
Test gain								
Mathematics		24.224	1.918	34.433	0.296	9.710	34.899	0.306
Verbal[b]		20.938	3.432	30.349	0.310	3.023	33.250	0.370
Reading		10.261	1.247	12.853	0.202	1.037	13.113	0.218
Vocabulary		10.677	2.195	12.772	0.164	1.981	14.257	0.251
Science		10.906	0.936	11.461	0.048	0.818	11.914	0.085
Writing[c]		9.999	1.696	15.306	0.347	1.673	16.073	0.378
Civics[c]		6.555	1.126	8.204	0.201	1.019	8.204	0.201

Source: Estimates by author based on data from the High School and Beyond study, 1980 and 1982.

[a]Estimates of the variance of measurement error for the sophomore and senior tests were obtained from Rock et al. (1985). They are based on Cronbach's coefficient alpha for formula scores. Estimates of the variance of measurement error for test gain were obtained by summing the sophomore and senior error variances; this procedure assumes that the measurement errors on the sophomore and senior tests are uncorrelated.

[b]The verbal test is the unweighted sum of the reading and vocabulary tests.

[c]These tests are presented here for the sake of completeness. They are not used in the analysis.

TABLE A-2
Average Course Enrollments in Grades Nine through Twelve by Graduation Status and Post-High School Plans

Course	High School Dropouts	High School Graduates/ Non-College-Bound	High School Graduates/ College-Bound	All Students
Vocational education				
Math-related voc.	0.432	0.909	0.700	0.763
Non-math-related	2.644	4.783	2.489	3.390
All vocational courses	3.076	5.692	3.189	4.153
Specific voc. math	0.112	0.128	0.051	0.085
Applied math	0.130	0.156	0.066	0.105
Mathematics				
Basic	0.238	0.136	0.051	0.097
General	0.789	0.633	0.279	0.451
Computer math	0.012	0.018	0.041	0.030
Prealgebra	0.163	0.234	0.230	0.227
Algebra 1	0.305	0.515	0.742	0.624
Geometry	0.115	0.255	0.709	0.493
Algebra 2	0.074	0.167	0.591	0.391
Precalculus	0.023	0.062	0.359	0.221
Calculus	0.000	0.006	0.093	0.053
All math courses	1.719	2.026	3.095	2.587
Science				
Survey	0.786	0.833	0.695	0.755
Biology	0.648	0.819	1.123	0.973
Chemistry	0.038	0.117	0.571	0.359
Physics	0.022	0.045	0.274	0.168
All science courses	1.494	1.814	2.663	2.255
English	3.018	3.924	4.075	3.945
Social studies	2.240	3.250	3.338	3.230
Fine arts	0.936	1.381	1.513	1.423
Foreign languages	0.285	0.459	1.579	1.057
Personal and other	2.396	2.836	3.004	2.898
Total credits	15.406	21.666	22.573	21.738
Sample size	855	3,759	6,347	10,961

TABLE A-3
Average Course Enrollments in Ninth and Tenth Grade Courses by Graduation Status and Post-High School Plans

Course	High School Dropouts	High School Graduates/ Non-College-Bound	High School Graduates/ College-Bound	All Students
Vocational education				
Math-related voc.	0.274	0.330	0.192	0.251
Non-math-related	1.566	1.613	0.972	1.261
All vocational courses	1.840	1.943	1.164	1.512
Specific voc. math	0.071	0.074	0.020	0.044
Applied math	0.077	0.067	0.023	0.044
Mathematics				
Basic	0.197	0.102	0.042	0.076
General	0.681	0.524	0.212	0.365
Computer math	0.005	0.004	0.005	0.005
Prealgebra	0.147	0.212	0.203	0.203
Algebra 1	0.269	0.453	0.664	0.555
Geometry	0.067	0.165	0.522	0.352
Algebra 2	0.037	0.072	0.178	0.127
Precalculus	0.011	0.014	0.042	0.029
Calculus	0.000	0.001	0.004	0.002
All math courses	1.414	1.547	1.872	1.714
Science				
Survey	0.677	0.734	0.631	0.674
Biology	0.520	0.653	0.842	0.747
Chemistry	0.018	0.024	0.094	0.062
Physics	0.011	0.010	0.014	0.012
All science courses	1.226	1.421	1.581	1.495
English	2.079	2.105	2.112	2.107
Social studies	1.366	1.424	1.433	1.425
Fine arts	0.677	0.733	0.789	0.760
Foreign languages	0.211	0.327	1.016	0.695
Personal and other	1.796	1.793	1.832	1.814
Total credits	10.757	11.434	11.842	11.610
Sample size	855	3,759	6,347	10,961

TABLE A-4

Average Course Enrollments in Eleventh and Twelfth Grade Courses by Predicted Sophomore Math Test Triptiles

Course	Lower Third	Middle Third	Upper Third	All Students
Vocational education				
Math-related voc.	0.341	0.617	0.561	0.509
Non-math-related	2.893	2.433	0.995	2.110
All vocational courses	3.234	3.050	1.556	2.619
Specific voc. math	0.060	0.046	0.012	0.040
Applied math	0.108	0.058	0.017	0.061
Mathematics				
Basic	0.053	0.007	0.002	0.020
General	0.143	0.070	0.041	0.084
Computer math	0.003	0.015	0.060	0.025
Prealgebra	0.030	0.031	0.012	0.024
Algebra 1	0.090	0.072	0.041	0.067
Geometry	0.048	0.208	0.160	0.139
Algebra 2	0.042	0.194	0.555	0.263
Precalculus	0.011	0.050	0.519	0.191
Calculus	0.000	0.003	0.150	0.050
All math courses	0.419	0.648	1.534	0.863
Science				
Survey	0.117	0.076	0.044	0.079
Biology	0.181	0.209	0.279	0.223
Chemistry	0.031	0.186	0.673	0.296
Physics	0.012	0.044	0.414	0.155
All science courses	0.341	0.515	1.410	0.753
English	1.732	1.803	1.949	1.828
Social studies	1.610	1.859	1.897	1.790
Fine arts	0.500	0.741	0.728	0.658
Foreign languages	0.110	0.285	0.681	0.359
Personal and other	0.959	1.123	1.143	1.076
Total credits	9.074	10.130	10.933	10.047
Sample size	3,648	3,647	3,666	10,961

TABLE A-5

Estimates of a Reduced-Form Sophomore Math Test Regression for All Students[a]

Right-Hand-Side Variable	Coeff.	Std. Error
Credits in eleventh/twelfth grade		
Basic math	-2.506	(0.469)
General math	-0.554	(0.221)
Computer math	2.267	(0.445)
Prealgebra	0.280	(0.419)
Algebra 1	0.042	(0.265)
Geometry	0.439	(0.191)
Algebra 2	1.529	(0.162)
Precalculus	2.221	(0.181)
Calculus	3.234	(0.300)
Specific vocational math	-0.840	(0.340)
Applied math	-0.819	(0.270)
Chemistry/physics	1.664	(0.121)
Biology/survey science	0.042	(0.127)
Math-related voc. ed.	0.376	(0.080)
Non-math-related voc. ed.	-0.217	(0.041)
English/social studies	-0.029	(0.053)
Foreign languages	0.780	(0.107)
Fine arts	0.227	(0.070)
Personal and other	-0.034	(0.068)
Graduation indicator	1.643	(0.304)
Female	-1.703	(0.135)
Black	-4.293	(0.222)
Asian	-1.729	(0.590)
Hispanic	-3.544	(0.207)
Urban	-0.003	(0.190)
Suburban	0.386	(0.147)
Northeast	-0.544	(0.200)
West	0.646	(0.213)
South	-1.113	(0.172)

(table continues)

TABLE A-5 (continued)

Right-Hand-Side Variable	Coeff.	Std. Error
Credits in ninth grade		
Basic math	-3.125	(0.372)
General math	-3.121	(0.272)
Computer math	-1.798	(1.712)
Prealgebra	-0.530	(0.300)
Algebra 1	1.426	(0.267)
Geometry	4.065	(0.389)
Advanced math	2.880	(0.430)
Specific vocational math	-1.948	(0.707)
Applied math	-1.344	(0.651)
Chemistry/physics	0.478	(0.797)
Biology	-0.320	(0.790)
Survey science	-0.793	(0.228)
Math-related voc. ed.	0.493	(0.219)
Non-math-related voc. ed.	-0.386	(0.112)
English/social studies	-0.351	(0.127)
Foreign languages	-0.009	(0.275)
Fine arts	0.187	(0.131)
Personal and other	-0.291	(0.125)
Math GPA in ninth grade	1.174	(0.073)
English GPA in ninth grade	0.393	(0.082)
Soc. studies GPA in ninth grade	0.162	(0.061)
For. lang. GPA in ninth grade	0.543	(0.098)
Science GPA in ninth grade	0.471	(0.079)
Other courses GPA in ninth grade	0.306	(0.089)
Expected postsecondary vocational education[b]	0.122	(0.099)
Constant	7.665	0.460
R-square	0.550	
Sample size	10,961	

[a]This equation is identical to the first-stage sophomore math test equations used throughout this paper.

[b]Years of postsecondary vocational education that the student expects to obtain.

TABLE A-6

First-Stage Estimates of the General 2SLS Model of Math Gain for Non-College-Bound Students[a]

Right-Hand-Side Variable	Sophomore Math Test Regression		Sophomore Science Test Regression		Sophomore Verbal Test Regression	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Credits in eleventh/twelfth grade						
Basic math	-2.032	(0.556)	-0.906	(0.311)	-1.305	(0.574)
General math	-1.409	(0.304)	-0.761	(0.170)	-1.721	(0.314)
Computer math	3.232	(0.961)	0.955	(0.539)	1.057	(0.993)
Prealgebra	-1.334	(0.707)	-0.126	(0.396)	-0.945	(0.730)
Algebra 1	0.237	(0.447)	0.294	(0.250)	0.441	(0.462)
Geometry	0.698	(0.370)	0.249	(0.207)	0.203	(0.382)
Algebra 2	1.565	(0.369)	0.790	(0.207)	1.601	(0.381)
Precalculus	1.291	(0.475)	0.009	(0.266)	-0.201	(0.491)
Calculus	4.511	(1.498)	1.540	(0.840)	2.979	(1.548)
Specific vocational math	-0.554	(0.451)	-0.030	(0.253)	0.254	(0.466)
Applied math	-0.620	(0.348)	-0.103	(0.195)	-0.358	(0.360)
Chemistry/physics	1.826	(0.306)	0.923	(0.172)	1.337	(0.316)
Biology/survey science	-0.053	(0.214)	0.287	(0.120)	0.322	(0.221)
Math-related voc. ed.	0.398	(0.112)	0.320	(0.063)	0.168	(0.116)
Non-math-related voc. ed.	-0.029	(0.055)	-0.002	(0.031)	-0.054	(0.056)
English/social studies	-0.043	(0.078)	0.024	(0.044)	0.205	(0.081)
Foreign languages	0.285	(0.271)	0.100	(0.152)	0.702	(0.280)
Fine arts	-0.031	(0.115)	0.106	(0.065)	0.030	(0.119)
Personal and other	-0.079	(0.106)	-0.116	(0.060)	-0.075	(0.110)
Graduation indicator	1.045	(0.373)	0.208	(0.209)	0.411	(0.386)
Female	-1.350	(0.205)	-1.601	(0.115)	-1.871	(0.212)
Black	-3.748	(0.346)	-2.930	(0.194)	-4.479	(0.357)
Asian	-2.674	(1.221)	-2.850	(0.684)	-4.786	(1.262)
Hispanic	-2.793	(0.279)	-2.250	(0.156)	-4.202	(0.288)
Urban	0.099	(0.288)	-0.435	(0.161)	0.214	(0.298)
Suburban	0.412	(0.218)	-0.080	(0.122)	0.153	(0.226)
Northeast	0.756	(0.307)	0.029	(0.172)	0.168	(0.317)
West	0.245	(0.337)	0.290	(0.189)	1.190	(0.348)
South	-0.870	(0.258)	-0.478	(0.144)	-0.803	(0.266)

(table continues)

TABLE A-6 (continued)

Right-Hand-Side Variable	Sophomore Math Test Regression Coeff.	Std. Error	Sophomore Science Test Regression Coeff.	Std. Error	Sophomore Verbal Test Regression Coeff.	Std. Error
Credits in ninth grade						
Basic math	-2.880	(0.491)	-0.402	(0.275)	-1.365	(0.507)
General math	-2.862	(0.378)	-0.293	(0.212)	-0.977	(0.390)
Computer math	-3.121	(2.913)	-1.936	(1.633)	0.530	(3.010)
Prealgebra	-0.229	(0.444)	0.288	(0.249)	0.095	(0.459)
Algebra 1	2.382	(0.400)	0.869	(0.224)	1.566	(0.413)
Geometry	4.214	(0.863)	1.037	(0.484)	2.416	(0.892)
Advanced math	3.081	(0.887)	1.159	(0.497)	3.522	(0.917)
Specific vocational math	-1.732	(0.874)	0.479	(0.490)	-1.668	(0.903)
Applied math	-1.340	(0.809)	0.073	(0.454)	-1.457	(0.836)
Chemistry/physics	2.779	(1.366)	-0.496	(0.765)	-1.246	(1.411)
Biology	-2.294	(1.376)	0.433	(0.771)	0.993	(1.422)
Survey science	-0.416	(0.319)	-0.303	(0.179)	-0.886	(0.330)
Math-related voc. ed.	0.242	(0.305)	0.679	(0.171)	0.474	(0.316)
Non-math-related voc. ed.	-0.302	(0.156)	-0.022	(0.088)	-0.336	(0.162)
English/social studies	-0.449	(0.176)	-0.414	(0.099)	-0.859	(0.182)
Foreign languages	-1.021	(0.501)	0.341	(0.281)	0.828	(0.518)
Fine arts	0.120	(0.199)	0.079	(0.112)	0.292	(0.206)
Personal and other	-0.086	(0.194)	-0.023	(0.109)	0.050	(0.200)
Math GPA in ninth grade	1.113	(0.110)	0.264	(0.061)	0.107	(0.113)
English GPA in ninth grade	0.288	(0.120)	0.065	(0.067)	0.741	(0.124)
Soc. studies GPA in ninth grade	0.108	(0.101)	0.112	(0.056)	0.493	(0.104)
For. lang. GPA in ninth grade	0.925	(0.198)	0.027	(0.111)	0.404	(0.204)
Science GPA in ninth grade	0.343	(0.124)	0.306	(0.069)	0.437	(0.128)
Other courses GPA in ninth grade	0.247	(0.134)	0.155	(0.075)	0.147	(0.138)
Expected postsecondary vocational education[b]	0.599	(0.115)	0.391	(0.065)	0.713	(0.119)
Constant	6.551	(0.605)	7.467	(0.339)	10.568	(0.625)
R-square	0.391		0.297		0.300	
Sample size	4,421		4,421		4,421	

[a]This equation is identical to the first-stage sophomore math test equations used throughout this paper.
[b]Years of postsecondary vocational education that the student expects to obtain.

TABLE A-7
Estimates of Alternative Biased Models of Determinants of Math Gain for Non-College-Bound Students

	Simple Difference Model with Omitted Soph. Math Test		Flooded Difference Model, Including 3 Soph. Tests with Measurement Error Ignored		Super Biased Senior Math Test Model, Excluding Proxies for Omitted Soph. Math Test	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Sophomore math test[a]	–	–	0.600	(0.013)	–	–
Sophomore science test[a]	–	–	0.143	(0.027)	–	–
Sophomore verbal test[a]	–	–	0.093	(0.014)	–	–
Credits in eleventh/twelfth grade						
Basic math	-0.376	(0.495)	-1.442	(0.456)	-4.041	(0.610)
General math	0.415	(0.269)	-0.175	(0.250)	-1.732	(0.332)
Computer math	-1.888	(0.870)	-0.774	(0.793)	1.564	(1.073)
Prealgebra	2.285	(0.637)	1.500	(0.582)	-0.007	(0.786)
Algebra 1	2.529	(0.396)	2.163	(0.365)	2.034	(0.488)
Geometry	1.922	(0.328)	2.472	(0.301)	3.972	(0.405)
Algebra 2	1.697	(0.327)	2.672	(0.301)	5.407	(0.403)
Precalculus	1.346	(0.426)	2.442	(0.389)	4.489	(0.525)
Calculus	1.197	(1.341)	3.623	(1.224)	9.718	(1.653)
Specific vocational math	0.071	(0.408)	-0.260	(0.372)	-0.789	(0.504)
Applied math	0.707	(0.311)	0.272	(0.287)	-0.457	(0.383)
Chemistry/physics	0.680	(0.275)	1.434	(0.253)	3.488	(0.339)
Biology/survey science	0.032	(0.191)	-0.133	(0.175)	-0.304	(0.235)
Math-related voc. ed.	0.161	(0.100)	0.388	(0.092)	0.988	(0.123)
Non-math-related voc. ed.	-0.119	(0.049)	-0.109	(0.045)	-0.082	(0.060)
English/social studies	0.086	(0.070)	0.051	(0.064)	0.012	(0.087)
Foreign languages	0.048	(0.242)	0.340	(0.221)	1.248	(0.298)
Fine arts	-0.149	(0.100)	-0.156	(0.092)	-0.101	(0.124)
Personal and other	-0.147	(0.094)	-0.109	(0.086)	-0.058	(0.116)
Graduation indicator	0.566	(0.331)	1.290	(0.305)	3.111	(0.409)
Female	-0.882	(0.178)	-0.758	(0.165)	-1.242	(0.220)
Black	0.080	(0.307)	-1.008	(0.293)	-5.097	(0.379)
Hispanic	0.007	(1.108)	-1.004	(0.237)	-4.159	(0.309)
Asian	-0.341	(0.250)	-0.569	(1.011)	-3.885	(1.366)
Suburban	-0.026	(0.195)	0.173	(0.178)	0.625	(0.241)
Urban	-0.053	(0.256)	0.072	(0.235)	0.199	(0.316)
Northeast	-0.588	(0.253)	-0.539	(0.231)	-0.524	(0.313)
West	0.333	(0.295)	0.336	(0.270)	0.706	(0.363)
South	-0.046	(0.222)	-0.415	(0.203)	-1.385	(0.273)
Constant	0.555	(0.349)	1.293	(0.363)	7.720	(0.430)
R-square	0.058		0.222		0.324	
Sample size	4,448		4,421		4,448	

[a]Endogenous variables subject to measurement error.

TABLE A-8

Estimate of the Prior Achievement Proxy Model for Non-College-Bound Students

Right-Hand-Side Variable	Coeff.	Std. Error
Credits in eleventh/twelfth grade		
Basic math	-2.517	(0.570)
General math	-1.018	(0.311)
Computer math	1.298	(0.992)
Prealgebra	0.897	(0.728)
Algebra 1	2.960	(0.456)
Geometry	2.647	(0.380)
Algebra 2	3.346	(0.381)
Precalculus	2.704	(0.491)
Calculus	6.039	(1.547)
Specific vocational math	-0.531	(0.465)
Applied math	0.004	(0.355)
Chemistry/physics	2.465	(0.316)
Biology/survey science	-0.049	(0.220)
Math-related voc. ed.	0.564	(0.115)
Non-math-related voc. ed.	-0.162	(0.056)
English/social studies	0.056	(0.080)
Foreign languages	0.315	(0.279)
Fine arts	-0.240	(0.118)
Personal and other	-0.205	(0.109)
Graduation indicator	1.720	(0.384)
Female	-2.182	(0.211)
Black	-3.781	(0.354)
Asian	-2.679	(1.261)
Hispanic	-3.166	(0.287)
Urban	0.095	(0.296)
Suburban	0.413	(0.225)
Northeast	0.180	(0.316)
West	0.544	(0.346)
South	-0.919	(0.265)

(table continues)

Table A-8 (continued)

Right-Hand-Side Variable	Coeff.	Std. Error
Credits in ninth grade		
Basic math	-2.471	(0.505)
General math	-2.424	(0.389)
Computer math	-5.621	(3.009)
Prealgebra	0.277	(0.457)
Algebra 1	2.693	(0.412)
Geometry	3.221	(0.891)
Advanced math	3.180	(0.916)
Specific vocational math	-1.866	(0.902)
Applied math	-1.149	(0.827)
Chemistry/physics	1.277	(1.410)
Biology	-0.741	(1.421)
Survey science	-0.850	(0.328)
Math-related voc. ed.	-0.277	(0.315)
Non-math-related voc. ed.	-0.013	(0.161)
English/social studies	-0.140	(0.182)
Foreign languages	-0.729	(0.515)
Fine arts	0.418	(0.205)
Personal and other	-0.014	(0.200)
Math GPA in ninth grade	0.961	(0.113)
English GPA in ninth grade	0.223	(0.124)
Soc. studies GPA in ninth grade	0.194	(0.104)
For. lang. GPA in ninth grade	0.709	(0.204)
Science GPA in ninth grade	0.539	(0.127)
Other courses GPA in ninth grade	0.155	(0.137)
Expected postsecondary vocational education	0.583	(0.119)
Constant	6.352	(0.622)
R-square	0.432	
Sample size	4,448	

TABLE A-9
Estimates of General 2SLS Model of Math Gain for Non-College-Bound Students, College-Bound Students, and All Students[a]

Right-Hand-Side Variable	Non-College-Bound Students Coeff.	Non-College-Bound Students Std. Error	College-Bound Students Coeff.	College-Bound Students Std. Error	All Students Coeff.	All Students Std. Error
Sophomore math test[a]	1.057	(0.071)	1.018	(0.068)	1.001	(0.053)
Sophomore science test[a]	-0.159	(0.219)	-0.127	(0.212)	-0.221	(0.164)
Sophomore verbal test[a]	0.076	(0.095)	0.058	(0.069)	0.085	(0.055)
Credits in eleventh/twelfth grade						
Basic math	-0.620	(0.504)	1.508	(0.802)	-0.058	(0.417)
General math	0.216	(0.286)	1.036	(0.284)	0.674	(0.198)
Computer math	-1.594	(0.870)	-1.261	(0.426)	-1.220	(0.392)
Prealgebra	2.199	(0.637)	1.577	(0.446)	1.883	(0.372)
Algebra 1	2.461	(0.402)	2.286	(0.284)	2.411	(0.236)
Geometry	2.034	(0.330)	1.304	(0.203)	1.492	(0.170)
Algebra 2	1.888	(0.347)	1.711	(0.162)	1.741	(0.144)
Precalculus	1.538	(0.439)	0.998	(0.206)	1.103	(0.181)
Calculus	1.635	(1.338)	0.419	(0.304)	0.649	(0.282)
Specific voc. math	0.575	(0.310)	1.612	(0.451)	0.664	(0.299)
Applied math	(0.001)	(0.405)	1.438	(0.385)	0.935	(0.239)
Chemistry/physics	0.884	(0.296)	0.499	(0.152)	0.669	(0.130)
Biology/survey science	0.031	(0.195)	-0.186	(0.134)	-0.087	(0.112)
Math-related voc. ed.	0.257	(0.115)	-0.048	(0.110)	0.132	(0.078)
Non-math-related voc. ed.	-0.115	(0.048)	-0.127	(0.065)	-0.120	(0.038)
English/social studies	0.084	(0.071)	-0.152	(0.065)	-0.035	(0.048)
Foreign languages	0.056	(0.250)	0.048	(0.110)	0.106	(0.103)
Fine arts	-0.134	(0.101)	-0.062	(0.076)	-0.045	(0.061)
Personal and other	-0.159	(0.095)	-0.159	(0.081)	-0.161	(0.062)
Graduation indicator	0.694	(0.335)	0.684	(0.505)	0.755	(0.268)
Female	-1.038	(0.291)	-0.503	(0.229)	-0.824	(0.188)
Black	-0.339	(0.492)	-0.223	(0.399)	-0.346	(0.309)
Hispanic	-0.560	(0.396)	-0.797	(0.616)	-0.738	(0.553)
Asian	-0.292	(1.163)	-0.364	(0.339)	-0.455	(0.256)
Suburban	-0.055	(0.199)	0.094	(0.174)	0.001	(0.133)
Urban	-0.123	(0.280)	0.200	(0.227)	-0.012	(0.179)
Northeast	-0.592	(0.250)	0.574	(0.219)	0.038	(0.163)
West	0.323	(0.308)	0.114	(0.227)	0.197	(0.183)
South	-0.158	(0.227)	-0.300	(0.198)	-0.234	(0.149)
Constant	1.420	(1.204)	1.144	(1.252)	1.418	(0.866)
R-square (T_2-T_1)	0.060		0.080		0.084	
Sample size	4,389		6,472		10,894	

[a]First-stage results for each group are reported in an appendix available from the author. The variables included in each first stage are identical to those listed in Appendix Table A-6.

[b]Endogenous variables subject to measurement error.

TABLE A-10
Estimates of General 2SLS Model of Math Gain for Each Predicted Sophomore Math Test Triptile[a]

Right-Hand-Side Variable	Lower Third Coeff.	Lower Third Std. Error	Middle Third Coeff.	Middle Third Std. Error	Upper Third Coeff.	Upper Third Std. Error
Sophomore math test[b]	0.975	(0.128)	0.839	(0.069)	0.848	(0.062)
Sophomore science test[b]	0.316	(0.213)	0.025	(0.220)	0.095	(0.195)
Sophomore verbal test[b]	-0.102	(0.098)	0.100	(0.089)	0.053	(0.062)
Credits in eleventh/twelfth grade						
Basic math	-0.321	(0.492)	2.994	(1.290)	2.926	(2.261)
General math	0.421	(0.301)	0.823	(0.344)	1.449	(0.404)
Computer math	-3.511	(2.338)	-1.221	(0.903)	-1.153	(0.373)
Prealgebra	1.552	(0.601)	1.484	(0.570)	2.251	(0.730)
Algebra 1	2.117	(0.374)	2.474	(0.401)	2.003	(0.418)
Geometry	1.111	(0.531)	1.122	(0.245)	1.881	(0.256)
Algebra 2	1.970	(0.555)	2.099	(0.262)	1.647	(0.169)
Precalculus	-0.426	(1.016)	1.832	(0.439)	1.387	(0.186)
Calculus	1.359	(7.107)	3.871	(1.843)	1.115	(0.271)
Specific voc. math	-0.099	(0.425)	1.079	(0.494)	2.279	(0.768)
Applied math	0.568	(0.315)	1.123	(0.446)	1.819	(0.697)
Chemistry/physics	0.249	(0.494)	0.862	(0.274)	0.508	(0.144)
Biology/survey science	0.082	(0.204)	-0.203	(0.202)	-0.200	(0.158)
Math-related voc. ed.	0.331	(0.157)	0.026	(0.123)	-0.098	(0.111)
Non-math-related voc. ed.	-0.169	(0.054)	-0.140	(0.069)	0.051	(0.087)
English/social studies	0.080	(0.077)	-0.115	(0.083)	-0.142	(0.089)
Foreign languages	0.226	(0.320)	-0.076	(0.182)	0.205	(0.119)
Fine arts	-0.083	(0.129)	-0.149	(0.098)	-0.033	(0.085)
Personal and other	-0.093	(0.118)	-0.152	(0.100)	-0.144	(0.090)
Graduation indicator	0.496	(0.341)	0.877	(0.585)	2.254	(1.363)
Female	-0.515	(0.284)	-0.729	(0.304)	-0.440	(0.232)
Black	-0.023	(0.428)	-0.031	(0.575)	0.328	(0.653)
Asian	-0.947	(1.522)	-0.559	(0.923)	-0.112	(0.649)
Hispanic	-0.003	(0.409)	-0.807	(0.415)	0.275	(0.482)
Urban	0.421	(0.312)	0.006	(0.298)	0.085	(0.267)
Suburban	0.248	(0.234)	-0.244	(0.231)	0.347	(0.190)
Northeast	0.200	(0.323)	-0.139	(0.288)	0.029	(0.220)
West	0.722	(0.406)	-0.128	(0.306)	0.254	(0.257)
South	0.102	(0.259)	-0.274	(0.260)	-0.536	(0.232)
Constant	-0.897	(1.180)	1.498	(1.430)	-0.006	(2.187)
R-square (T_2-T_1)	0.043		0.076		0.080	
Sample size	3,610		3,632		3,650	

[a]First-stage results for each triptile are reported in an appendix available from the author. The variables included in each first stage are identical to those listed in Appendix Table A-6. Triptile samples are described in a note in Appendix Table A-4.

[b]Endogenous variables subject to measurement error.

TABLE A-11
Estimates of General 2SLS Model of Math Gain for Students with Predicted Sophomore Math Scores between the 50th and 90th Percentiles[a]

Right-Hand-Side Variable	Coefficient	Standard Error
Sophomore math test[b]	0.941	(0.066)
Sophomore science test[b]	-0.124	(0.202)
Sophomore verbal test[b]	0.086	(0.068)
Credits in eleventh/twelfth grade		
Basic math	3.468	(1.373)
General math	1.666	(0.387)
Computer math	-1.282	(0.554)
Prealgebra	3.156	(0.633)
Algebra 1	2.470	(0.373)
Geometry	1.536	(0.215)
Algebra 2	2.050	(0.172)
Precalculus	1.675	(0.221)
Calculus	1.564	(0.580)
Specific vocational math	1.100	(0.585)
Applied math	1.259	(0.554)
Chemistry/physics	0.921	(0.174)
Biology/survey science	-0.255	(0.161)
Math-related voc. ed.	0.110	(0.105)
Non-math-related voc. ed.	-0.027	(0.070)
English/social studies	-0.080	(0.080)
Foreign languages	0.061	(0.113)
Fine arts	-0.047	(0.082)
Personal and other	-0.262	(0.095)
Graduation indicator	0.565	(0.817)
Female	-0.862	(0.269)
Black	-0.379	(0.558)
Asian	-0.449	(0.709)
Hispanic	-0.544	(0.423)
Urban	0.335	(0.262)
Suburban	0.011	(0.149)
Northeast	0.177	(0.236)
West	0.383	(0.254)
South	-0.403	(0.227)
Constant	11.280	(1.498)
R-square (T_2-T_1)	0.111	
Sample size	4,368	

[a]First-stage results are reported in an appendix available from the author. The variables included in the first stage are identical to those listed in Appendix Table A-6.
[b]Endogenous variables subject to measurement error.

Endnotes

[1]For college-bound students, an additional two years of foreign language was recommended.

[2]The three-year social studies requirement was implemented in twenty-seven states, while four states implemented the 1/2-year computer science requirement (Education Commission of the States, 1987).

[3]Hanushek (1986) reviewed 147 different educational production function studies published since the Coleman Report in 1966. As indicated in Table 8 of his paper, of the 112 studies that estimated the effect of student/teacher ratios on test performance, only 9 reported statistically significant, positive coefficients, while 14 reported statistically significant, negative coefficients. Similar results were observed for teacher education, teacher salaries, per-pupil expenditures, and teacher experience, although in the latter case, statistically significant, positive coefficients were reported in 33 out of 109 studies. See also Hanushek (1981).

[4]Teacher quality may be considered context specific if, for example, the types of skills best suited to teaching an applied mathematics course to low-achieving students differ from the skills best suited to teaching a trigonometry course to college-bound students. As pointed out by Hanushek (1986), if teacher quality is largely context specific, rather than universal, standard educational production functions will be unable to capture the sources of teacher productivity.

[5]Critiques of the original Coleman, Hoffer, and Kilgore analysis (together with their replies) can be found in the Harvard Education Review (November 1981) and in two issues of Sociology of Education (April/July 1982 and October 1983). Greeley (1982) conducted a similar analysis that examined the differential effects of Catholic and public schools on black and Hispanic students. Noell (1982) and Murnane, Newstead, and Olsen (1985) also conducted analyses using only the senior cohort. They attempted to control for unobserved differences in public and private school students using econometric techniques designed to correct for selection bias.

[6]These statistical methods are evaluated in Section IV.

[7]Alexander and Pallas (1985) point out that the estimated private school advantage is small relative to the accumulated variation in test scores up through sophomore year--less than 0.1 standard deviation. This comparison is misleading, however, since the estimated private school advantage pertains to the growth in test scores over two years. Over this interval, Hoffer, Greeley, and Coleman (1985) estimate that Catholic school students learn the equivalent of three grades' worth of material, compared to two grades' worth for public school students, on average.

[8]Hoffer, Greeley, and Coleman (1985) report (see their Table 2.8) that Catholic school students take an average of 2.74 advanced math courses, compared to 2.08 courses taken by public school students. The authors argue that the higher advanced math enrollments in private school are, in fact, an outcome of private school attendance. Goldberger and Cain (1982), in their critique of the earlier Coleman, Hoffer, and Kilgore (1982) report, argue the opposite--that enrollment and high school track decisions are determined primarily, if not exclusively, by nonschool factors such as unmeasured student ability and family background. This is clearly an important area for further research.

[9]One exception is the research in the First International Mathematics Study, also discussed in Husen (1972).

[10]The authors analyzed the effects of total semesters of coursework in algebra, advanced algebra, geometry, trigonometry, and calculus. Similar results were obtained, they report, when business, general math, prealgebra, and other math courses were also included in the total.

[11]Fetters, Stowe, and Owings (1984) compared transcript and self-reported course enrollments, using the HS&B sophomore data base. Assuming that the transcript data were correct, they found that self-reported enrollments in mathematics tended to be somewhat overstated with a reliability of only 70 percent.

[12]In fact, the same authors in a paper cited earlier (Alexander and Pallas, 1985) criticized Hoffer, Greeley, and Coleman (1985) for failing to correct for measurement error in the sophomore (HS&B) tests. Indeed, estimates in the two papers differed, in large part, because Alexander and Pallas (1985) used a statistical technique designed to correct for measurement error in the sophomore tests while Hoffer, Greeley, and Coleman (1985) ignored the problem. Section IV presents a detailed analysis of alternative estimators designed to correct for measurement error.

[13]See Alexander and Pallas (1984) for an analysis of the effectiveness of alternative course patterns.

[14]The proportionate increase in the variance of parameter estimates due to measurement error in ΔT_i is given by reciprocal (1 - unreliability rate/(1-R^2)) = reciprocal (0.2391), given that $R^2 = 0.08$ and the unreliability rate is 0.70.

[15]Robin S. Horn, Mark Braddock, and I prepared the classification of math and non-math-related vocational courses for the National Assessment of Vocational Education. Becky Hayward and Nancy Adelman provided helpful comments.

[16]All estimates reported in this paper were computed using the HS&B transcript sample weight, TRWT.

[17]A one Carnegie credit course typically meets for five fifty- to fifty-five-minute periods per week for an entire school year.

[18]Failed courses, which represented less than 5 percent of all enrollments, were included because the empirical evidence suggested that these courses made some contribution to growth in math test scores.

[19]Rock et al. (1985) found that the HS&B science test was highly correlated with both their mathematics and verbal factors. In a model that includes the HS&B math, composite verbal, and

science tests, it may be reasonable to interpret the science test as a measure of math skills not explicitly measured in the mathematics test and/or a measure of mathematics aptitude.

[20]As explained in Section VI, this was done so as to avoid introducing bias, due to sample truncation, into the 2SLS estimates presented in that section.

[21]All equation systems of this type will include an equation similar to (4.1) unless all of the prior determinants of achievement, external to the family, are measured.

[22]The fact that high school spans four years is therefore a poor reason to study the gain in achievement over a four-year period, rather than over four consecutive one-year periods.

[23]The papers discussed earlier by Welch, Anderson, and Harris (1982), Schmidt (1983), and Coleman, Hoffer, and Kilgore (1982) rely on the prior achievement proxy model. As already mentioned, this probably accounts for the inflated coefficients in mathematics courses in the first two studies.

[24]This conclusion is also based on the implicit assumption that enrollment decisions are influenced by annually (or more frequently) updated student knowledge of their own achievement levels.

[25]I argued earlier that measurement error in their prior test score also might account for the estimated pattern of coefficients.

[26]The parameter m is always ≥ 0 and ≤ 1, since the largest value of R^2 is $1-\lambda$ and the minimum value of λ and R^2 is zero.

[27]As indicated in equations (4.10) and (4.16), the expected coefficient on course enrollments (X_2) in the prior achievement proxy model is $(\alpha_2 + \theta_2\pi_2)$ and in the model with a fallible measure of T_1 is $(\alpha_2 + m\theta_2\eta_2)$. The former model exhibits less bias if $\pi_2 < m\eta_2$. Define C as the matrix of coefficients from the regression of vector X_1 on vector x_2, given Z. Then η_2 can be related to π_2 by considering the consequences of dropping X_1 from the model: $\eta_2 = \pi_2 + C\pi_1$. The prior achievement proxy model exhibits lower bias if $\pi_2 < m(\pi_2 + C\pi_1)$ or, equivalently, if $\pi_2 <$

$C\pi_1 m/(1-m)$. Thus the prior achievement proxy model exhibits lower bias if X_1 and X_2 are highly correlated (as reflected in C) and if measurement error is high (as reflected in m).

[28]In the latter case, if errors in measurement are correlated, it will also be necessary to adjust for inflated covariances among the variables measured with error. Often, however, it is not unreasonable to assume that measurement errors are uncorrelated, which eliminates the need to obtain estimates of the correlation structure of errors.

[29]See Allen and Yen (1979).

[30]This follows because $\sigma_v^2 = \mathrm{Var}(T) - \mathrm{Cov}(T, T') = \mathrm{Var}(T^* + \upsilon_1) - \mathrm{Cov}(T^* + \upsilon_1, T^* + \upsilon_2) = + \sigma_v^2$ where T^* = true test performance, and υ_1 and υ_2 are independent measurement errors with equal variance σ_v^2. Although sampling variation could cause small differences between the variances of T and T′, large differences could suggest that the two tests are measuring different things.

[31]Since average prior achievement is higher among private school students than public school students, upward bias in the coefficient on prior achievement (due to overcorrection for measurement error) would cause downward bias in the estimated private school effect. Jencks (1985) also criticizes Alexander and Pallas (1985) for including only one prior achievement score on the right-hand side of their achievement equation.

[32]Since tests were administered to the students in the HS&B data during the spring of their sophomore and senior years, coursework taken during the tenth grade could arguably make a small contribution to the growth in achievement between tests, contrary to the assumption made in the text. As a precautionary step, therefore, I do not actually use tenth grade enrollments in predicting T_1. This step ensures that the estimates are consistent, with only a small loss in efficiency.

[33]Additional exclusion restrictions on the elements of Z that belong in equation (4.5) could, if correct, also help provide independent variation in T_1, thereby increasing the efficiency of the estimates.

[34]This characteristic sharply differentiates the high school experience from the typical elementary school experience, in which students are taught predominantly by a single teacher. See Hanushek (1971) for an empirical study of the effectiveness of elementary school teachers.

[35]The so-called residual growth estimator is an ad hoc two-step estimator that, in my view, has no theoretical justification. In step one, T_2 is regressed on T_1. In step two, the residual from this equation, say $T_2 - \hat{\rho}_0 - \hat{\rho}T_1$, is used as the dependent variable in an equation that includes Z *and* X_2. This estimator is exactly equivalent to imposing the value $\theta_2 = \rho$ in estimating equation (4.5). In practice, this estimator may outperform the estimator with uncorrected measurement error (#4) because: (1) the downward bias in $\hat{\rho}$ due to measurement error in T_1 is not exacerbated by multicollinearity (as in estimator #4) and (2) the exclusion of Z and X_2, variables that are positively correlated with T_1, creates a positive bias in $\hat{\rho}$. There is absolutely no guarantee, however, that the negative and positive biases in ρ (as a surrogate estimate of θ_2) exactly offset each other.

[36]Of course, in the unlikely event that the true value of θ_2 were somehow known, a least squares estimator that imposed the true value of θ_2 would generate more efficient estimates than the 2SLS estimator. In addition, if θ_2 were "approximately known" (that is, known with a reasonably high degree of precision), an estimator that imposed the approximately known value of $\hat{\theta}_2$ could generate parameter estimates with lower mean squared error than the 2SLS estimator. In effect, a small amount of bias would be deliberately introduced in return for greater precision. Unfortunately, our ex ante uncertainty about θ_2 is rather large. In fact, our consistent estimate of θ_2 differs radically from the biased estimates of θ_2 produced in most previous studies.

[37]The prior achievement proxy model does not directly fit within the "family" of models discussed in the text because it includes an additional vector of right-hand-side variables: X_1. However, an equivalent value (or values) of $\tilde{\theta}$ is implicitly defined by equating the biased coefficient on X_2 (say, a) with its theoretical value from equation (5.1): $\alpha_2 + (\theta_2 - \tilde{\theta})\eta_2$. Then, $\tilde{\theta} = \theta_2 - (a - \alpha_2)/\eta_2$.

[38]The first-stage equation includes the following instrumental variables: eleventh/twelfth grade enrollments (X_2), ninth grade enrollments, demographic characteristics (Z), and ninth grade gradepoint averages (GPAs) in math, English, social studies, foreign language, science, and other courses, and a measure of postsecondary educational expectations. Similar, but less precise, estimates were obtained with the GPA and expectations variables removed.

[39]This conclusion is based on the assumption that all mathematics learning is school based. If some learning of mathematics takes place in nonschool experiences, and the incidence of nonschool mathematics learning is related to math ability, the coefficient on prior math achievement (θ_2) reflects both the negative depreciation of skills and the positive contribution of nonschool learning that is predicted by T_1.

[40]These formulas are based on the assumption that math learning depends on Z, X_2, and prior math achievement. Empirical results presented later in this section indicate that this hypothesis cannot be rejected. More complicated formulas could be used to retrieve estimates of λ if additional tests (measured with error) were included in the model.

[41]More fundamentally, this study points to the need to collect and validate more extensive longitudinal data on mathematics proficiency, tied to extensive school and student-level data and high school transcript data. From a psychometric standpoint, the models presented in this study provide a new and potentially powerful mechanism for validating alternative tests of mathematics achievement and aptitude.

[42]R^2 as a percentage of the usable variance of $(T_2 - T_1)$ is given by reported R^2 divided by the reliability of $(T_2 - T_1)$ (0.061/0.296 = 0.206).

[43]Fetters, Stowe, and Owings (1984) report that the reliability of the mother's education variable is over 80 percent. I did not include student-reported parental income as an explanatory variable because of its low reliability, 50 percent, as estimated by Fetters and his colleagues.

[44]The only major difference occurs in the estimated calculus coefficients: 1.64 in both 2SLS models versus 1.94 in both EV models. Given that the standard error on each of these coefficients is about 1.34, these differences are obviously not statistically significant.

[45]These empirical results were drawn from Tables 6, 8, A-7, and A-8.

[46]As indicated in endnote 37, an illustrative estimate of $\tilde{\theta}$ is given by

$$\tilde{\theta}_2 = \hat{\theta}_2 - (\hat{a} - \hat{\alpha}_2)/\hat{\eta}_2$$

$$= 0.947 - (6.039 - 1.647)/8.5$$

$$= 0.430,$$

where $\hat{a}$ is an estimated coefficient from the prior achievement proxy model (I have arbitrarily used the calculus coefficient), $\hat{\theta}_2$ and $\hat{\alpha}_2$ (calculus) are consistent 2SLS estimates, and $\hat{\eta}_2$ is the calculus coefficient from the auxiliary T_1 regression in Table 7.

[47]See Hausman and Wise (1977) or Maddala (1983) for a discussion of truncation bias. The estimates used in predicting sophomore math achievement are presented in Appendix Table A-5.

[48]This prediction, which cannot be derived from Table 11, is based on a model that allows for and finds modest diminishing returns in the effectiveness of math-related vocational education.

[49]The actual gain of 0.85 points is reported in a table that is not in the present paper.

References

Alexander, Karl L., and Aaron M. Pallas. "Curriculum Reform and School Performance: An Evaluation of the New Basics." American Journal of Education 92 (1984): 391-420.

Alexander, Karl L., and Aaron M. Pallas. "School Sector and Cognitive Performance: When Is a Little a Little?" Sociology of Education 58 (1985): 115-128.

Allen, Mary J., and Wendy M. Yen. Introduction to Measurement Theory. Monterey, Calif.: Brooks/Cole Publishing Company, 1979.

Bennett, William J. American Education: Making It Work (report prepared for the president and the American people). Washington, D.C.: Department of Education, 1988.

Boardman, Anthony E., and Richard J. Murnane. "Using Panel Data to Improve Estimates of the Determinants of Educational Achievement." Sociology of Education 52 (1979): 113-121.

Brown, Cynthia, et al. "The Secondary Schools Taxonomy." Discussion Paper for the National Assessment of Vocational Education, U.S. Department of Education, Washington, D.C. February, 1989.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. "Cognitive Outcomes in Public and Private Schools." Sociology of Education 55 (1982): 65-76.

Coleman, James S., et al. Equality of Educational Opportunity. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Government Printing Office, 1966.

Congressional Budget Office, the Congress of the United States. Educational Achievement: Explanations and Implications of Recent Trends. Washington, D.C.: Congressional Budget Office, 1987.

Congressional Budget Office, the Congress of the United States. Trends in Educational Achievement. Washington, D.C.: Congressional Budget Office, 1986.

Education Commission of the States. Minimum High School Graduation Requirements in 1980, 1985, and 1987: Standard Diplomas. Denver, Colo.: Education Commission of the States, 1987.

Fetters, William B., Peter S. Stowe, and Jeffrey A. Owings. High School and Beyond, a National Longitudinal Study for the 1980s: Quality of Responses of High School Students to Questionnaire Items. Washington, D.C.: National Center for Education Statistics, 1984.

Fuller, Wayne A. Measurement Error Models. New York: John Wiley and Sons, 1987.

Goldberger, Arthur S., and Glen G. Cain. "The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report." Sociology of Education 55 (1982): 103-122.

Greeley, A. Catholic High Schools and Minority Students. New Brunswick, N.J.: Transaction, 1982.

Hallinan, Maureen T. (ed.). Sociology of Education. Washington, D.C.: American Sociological Association, 1982.

Hanushek, Eric. "The Economics of Schooling." Journal of Economic Literature 24 (1986).

Hanushek, Eric. "Throwing Money at Schools." Journal of Policy Analysis and Management 1 (1981): 19-41.

Hanushek, Eric. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data". American Economic Review Proceedings 61 (1971): 280-88.

Harvard Education Review. November, 1981. (Entire journal.)

Hausman, Jerry A., and David A. Wise. "Social Experimentation, Truncated Distributions, and Efficient Estimation." Econometrica 45 (1977): 919-939.

Hoffer, Thomas, Andrew M. Greeley, and James S. Coleman. "Achievement Growth in Public and Catholic Schools." Sociology of Education 58 (1985): 74-97.

Husen, T. "Does More Time in School Make a Difference?" Saturday Review, April 29, 1972, 32-35.

Jencks, Christopher. "How Much Do High School Students Learn?" Sociology of Education 58 (1985): 128-135.

Koretz, Daniel M., personal communication with author, 1988.

Lord, Frederic M., and Melvin R. Novick. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison-Wesley Publishing Company, 1968.

Maddala, G. S. Limited-Dependent and Qualitative Variables in Economics. Cambridge, Mass.: Cambridge University Press, 1983.

McKnight, Curtis, et al. The Underachieving Curriculum: A National Report on the Second International Mathematics Study. Champaign, Ill.: Stipes Publishing Company, 1987.

Meyer, Robert. "Beyond Academic Reform: The Case for Integrated Vocational and Academic Education." Mimeograph, Institute for Research on Poverty, University of Wisconsin at Madison, 1990.

Murnane, Richard J., Stuart Newstead, and Randall J. Olsen. "Comparing Public and Private Schools: The Puzzling Role of Selectivity Bias." Journal of Business and Economic Statistics 3 (1985): 23-35.

Murnane, Richard J. The Impact of School Resources on the Learning of Inner City Children. Cambridge, Mass.: Ballinger, 1975.

National Assessment of Vocational Education. First Interim Report from the National Assessment of Vocational Education. Washington, D.C.: Department of Education, 1988.

National Commission on Excellence in Education. A Nation at Risk: The Imperative for Educational Reform. Washington, D.C.: Department of Education, 1983.

National Science Board Commission on Precollege Education in Mathematics, Science and Technology. Educating Americans for the 21st Century: A Plan of Action for Improving Mathematics, Science and Technology Education for All American Elementary and Secondary Students so that Their Achievement Is the Best in the World by 1995 (report to the American people and the National Science Board). Washington, D.C.: 1983.

Noell, Jay. Public and Catholic Schools: A Reanalysis of "Public and Private Schools." Sociology of Education 55 (1982): 123-132.

Pallas, Aaron M., and Karl L. Alexander. "Sex Differences in Qualitative SAT Performance: New Evidence on the Differential Coursework Hypothesis." American Educational Research Journal 20 (1983): 165-182.

Rock, Donald A., et al. "Psychometric Analysis of the NLS and High School and Beyond Test Batteries." Educational Testing Service. Prepared for the National Center of Education Statistics, U.S. Department of Education, September 1985.

Schmidt, William H. "High School Course-Taking: Its Relationship to Achievement." Journal of Curriculum Studies 15 (1983): 311-332.

Welch, Wayne W., Ronald E. Anderson, and Linda J. Harris. "The Effects of Schooling on Mathematics Achievement." American Educational Research Journal 19 (1982): 145-153.