DOCUMENT RESUME

ED 395 968 TM 025 106

AUTHOR	Fincher, Cameron
TITLE	Trends in the Predictive Validity of the Scholastic Aptitude Test.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-90-13
PUB DATE	Dec 90
NOTE	105p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Ability; Academic Achievement; *Change; College Entrance Examinations; College Freshmen; Correlation; *Educational Trends; *Grade Point Average; Higher Education; High Schools; High School Students; *Predictive Validity; *Test Use
IDENTIFIERS	*Scholastic Aptitude Test; *University of Georgia

ABSTRACT

Changes in the predictive validity of the Scholastic Aptitude Test (SAT) that have been observed in the University System of Georgia (USGA) over the past two decades were studied. Several analyses were performed using student data from 27 USGA schools to determine the effectiveness of the SAT in predicting freshman grades. National trends indicated that the validity coefficients for high school record (HSR) and HSR+SAT declined from 1964 to 1982. Findings from Georgia also demonstrate an appreciable decrease in validity coefficients for the past 6 or 7 years. The decrease in SAT validity coefficients is accompanied by similar variations for high school average (HSA) coefficients and for multiple correlation coefficients involving combinations of HSA and SAT. The same general trends were observed when annual variations were smoothed by computing 5- and 10-year averages. These general trends were affected by the variability of student abilities, as measured by the SAT, and academic performance, as reflected in high school and college grades. It is cautioned that decreases in the magnitude of validity coefficients do not mean a decline in the meaning and significance of the SAT as a measure of verbal and mathematical abilities. Appendix A contains 5-year averages in validity coefficients, and Appendix B contains profiles of cooperating institutions. (Contains 9 figures, 15 tables, 27 figures in Appendix A, 23 figures in Appendix B, and 7 references.) (SLD)

RR-90-13

RESEARCH REPORT

Trends in the Predictive Validity of the Scholastic Aptitude Test

Cameron Fincher

Educational Testing Service
Princeton, New Jersey
December 1990

TRENDS IN THE PREDICTIVE VALIDITY OF THE SCHOLASTIC APTITUDE TEST

by

Cameron Fincher

Institute of Higher Education

The University of Georgia

FOREWORD

This study of SAT validity trends could not be conducted quickly or easily. When asked to do the study, I was confident that the great majority of needed data was already on computer tape and merely a matter of pulling and analyzing computer files. I was also confident that a survey form could be quickly developed, printed, mailed, and completed by cooperative colleagues who would respond to my personal request for information and assistance.

I was completely wrong about the former and only partially right about the latter. It was necessary to re-code massive amounts of the University System of Georgia data that has been distributed annually since 1958. It was also necessary to harness the capabilities of new computer programs at a time when other demands on staff time interfered at every conceivable opportunity.

Despite distractions and unforgiveable delays in completing this report, I am indebted to many others for whatever information it might contain. Professional colleagues on other campuses did respond generously, but demands on their time were similar to demands on mine. And unfortunately, many of them simply did not have longitudinal data that would serve survey purposes. Several gave the study a strong endorsement while phoning or writing about their inability to cooperate.

Special words of praise are due Mrs. Joyce Placek and Mrs. Susan Sheffield who generated the tables and graphs we needed. Words of appreciation are also due Pete Foley, Carolyn Griswold, and Wes Wicker (graduate assistants in the Institute) who retrieved voluminous data and forced them into columns and rows that made sense. If any of us had hopes that sophisticated computer technology would lighten our clerical loads, we were obviously disappointed.

In brief, the following (revised) report has been much delayed <u>but</u> prepared in a hurry! I am now hopeful that the report will be informative <u>and</u> useful to the professional staff of ETS who requested and funded the study.

Cameron Fincher
May 22, 1990

CONTENTS

LIST OF TABLES

Tables (continued)

LIST OF FIGURES

APPENDIX A

Appendix A (continued)

Appendix A (continued)

APPENDIX B

Appendix B (continued)

TRENDS IN THE PREDICTIVE VALIDITY OF THE SCHOLASTIC APTITUDE TEST

Throughout the 1970s as mean scores on the SAT declined nationally, there were expressions of confidence that the decline in average performance did not affect the predictive validity of SAT verbal and mathematical scores. When the range of scores and grades was not restricted appreciably, the correlation between scores and grades remained relatively stable. When SAT-V and SAT-M were combined with high school average (HSA), multiple correlations averaged about .58 (Donlon, 1984).

More recently, criticisms of the SAT's predictive validity have implied that multiple correlations of HSA+SAT with college grades is sustained predominantly by the higher coefficients derived for HSA with freshman GPA. With access to higher education by a more culturally diverse student population, correlations between the verbal and mathematical abilities measured by the SAT and GPA have supposedly declined while correlations between HSA and GPA have increased. Thus, a decline in the SAT's predictive validity presumably followed the decline in mean scores, and critics of the SAT advocate either its discontinuance or its modification to reflect more accurately the academic achievement of entering college freshmen.

The objective of this study is to summarize the changes in the SAT's predictive validity that have been observed in the University System of Georgia (USGA) and other similar institutions over the past two decades. The University System of Georgia is a valuable source of such information because predictive coefficients have been computed for each institution since 1957, the year in which the SAT was first adopted as a systemwide entrance requirement. Similar institutions have been selected on an informal basis and asked to submit, if available, comparable information on the predictive usefulness of the SAT. The specific objective of the study, however, is not to compare institutional data <u>but</u> to detect changes in such data over a period of time and to identify trends where they can be described with reasonable accuracy.

The overall design and procedures of the study have been influenced by the author's experience in studying the effectiveness of the SAT over a thirty-year period. The first of four related studies was conducted for the years 1958-1970 and analyzed validity coefficients for thirteen years in the University System of Georgia. The three other studies were sponsored by the College Board and dealt with (1) the SAT's predictive validity for adult learners, (2) the SAT's usefulness in advisement, counseling, placement, and other academic decisions, and (3) the predictive contribution of the SAT when used in conjunction with high school grades. With information published in the College Board's Technical Handbook for the SAT, these USGA normative data, published annually since 1958, constitute the baseline for this study of recent changes.

The University System of Georgia (in 1989) has a total of 34 institutions, but for the purposes of this study only 27 have been used. The Medical College of Georgia has very few undergraduate students and has been excluded. Six other institutions were excluded because institutional data were not conducive to the analysis of trends.

The present study of the SAT's effectiveness in predicting freshman grades has been conducted in several stages which may be briefly described as follows:

1. The computation of five-year averages for the validity coefficients of SAT-V, SAT-M, HSA, and SAT+HSA over a thirty-year period for 27 institutions within the University System of Georgia. Each set of validity coefficients has been plotted for the 27 institutions (as a total group) and for each of the USGA units as a separate institution of higher education. A classification of USGA institutions, as grouped for purposes of this study, is given in Table I.

2. The computation and comparison of five-year averages for the standard deviations of SAT-V, SAT-M, and FGPA in the units (n=18) of the University System <u>and</u> for HSA and FGPA for the same set of institutions.

3. The computation of ten-year averages for the means and standard deviations of the predictor variables, plus the regression slopes of their respective validity coefficients across each ten-year period. These computations have been tabled by institution for the 18 institutions that were units of the University System throughout the years of 1958-1987.

4. The analysis of variance in predictive validity by gender by decade (1970s vs 1980s) for SAT-V, SAT-M, HSA, and SAT+HSA within University System institutions, as classified and grouped in Table 1 for purposes of this study.

5. The analysis of variance in predictive validity by decade by level of variability (as suggested by mean standard deviations) for units (n=27) of the University System.

6. The analysis of trends in validity coefficients for SAT scores and high school grades (HSA or HSR) for cooperating institutions who provided data and information in response to a survey of colleges and universities likely to have longitudinal data on the SAT's predictive efficiency.

7. The summary of informal information received from responding institutions who could not provide longitudinal data (e.g., some institutions did not conduct periodic or longitudinal studies and volunteered their reasons for not doing so).

Table 1. Classification of Institutions within the University System of Georgia (1958-1987)

Category	Institutions
UNIVERSITIES (n=3)	University of Georgia (UGA) Georgia Institute of Technology (TECH) Georgia State University (GSU)
REGIONAL SENIOR INSTITUTIONS (n=5)	Georgia Southern College (GSOU) Georgia College (GAC) North Georgia College (NGA) West Georgia College (WGA) Valdosta State College (VSC)
SENIOR COLLEGES (n=7)	Armstrong State College (ARSC) Augusta College (AUGC) Columbus College (COLC) Georgia Southwestern (GASW) Kennesaw State College (KNSC) Clayton State College (CLSC) Southern College of Technology (SCT)
RESIDENTIAL TWO-YEAR (n=3)	Abraham Baldwin College (ABAC) Middle Georgia College (MDGA) South Georgia College (SOGA)
HISTORICALLY BLACK INSTITUTIONS (n=3)	Albany State College (ALSC) Fort Valley State College (FVSC) Savannah State College (SSC)
OTHER TWO-YEAR COLLEGES (n=6)	Albany Junior College (ALJC) Brunswick Junior College (BJC) Dalton Junior College (DJC) Floyd Junior College (FLJC) Gainesville Junior College (GJC) Macon Junior College (MJC)

Establishing a Baseline

As a baseline by which to assess changes in validity coefficients for SAT-V, SAT-M, HSA, and SAT+HSA, summaries of previous studies are provided in Table 2 and Figure 1. The validity coefficients reported in Table 2 are re-tabled from the College Board Validity Study Service reported in the 1984 Technical Handbook and from the meta-analysis reported in 1985 on the USGA "central prediction model" (Fincher, 1986). For male students, Table 2 shows average correlation coefficients of +.51 (HSA+SAT) nationally and +.58 (HSA+SAT) for freshmen in the University System of Georgia. Thus in the CEEB/VSS data, when SAT scores are combined with HSA, they contribute an additional seven percentage points to the variance accounted for by high school grades alone. In the USGA data the SAT contribution raises the multiple correlation coefficient to +.58, a gain of ten correlation points over the +.48 computed for HSA.

For women students, the use of SAT scores in CEEB/VSS data raises the multiple correlation coefficient from +.49 to +.57 and accounts for an additional eight percent in the explained variance. In USGA data, SAT scores raise the multiple correlation to +.63, a gain of nine coefficient points over HSA used alone. In brief, the gains in predictive efficiency (by using the SAT) tend to be higher in USGA than in national validity studies.

National trends in average validity coefficients have been plotted from data reported in the CEEB Technical Handbook (Chapter VII). As shown in Figure 1, these trends reveal a decline from +.59 to +.54 for HSR+SAT (over a nineteen-year period). HSR has declined from +.52 to +.46 while SAT-V has dropped two coefficient points and SAT-M has gained five coefficient points. In other words, the slight loss of two coefficient points for SAT-V (in predictive validity) is not as significant as the loss of six points for HSR and the loss of five points for HSR+SAT.

In summarizing national trends, we may conclude that validity coefficients for HSR and HSR+SAT declined five or six points from 1964 to 1982. SAT-V was fairly stable in its predictive efficiency, losing only two coefficient points, and SAT-M apparently gained in predictive efficiency. These observed changes in average validity coefficients should be of assistance in assessing trends in more recent years.

Trends in Five-Year Averages

Simple five-year averages for the validity coefficients of SAT-V, SAT-M, HSA, and SAT+HSA are given in Figure 2. The averages have been computed across institutions (n=18) and plotted on the mid-year of each five-year period. Changes in the correlation of SAT and HSA with FGPA are displayed for the thirty years in which the SAT has been required for entrance to units of the University System of Georgia.

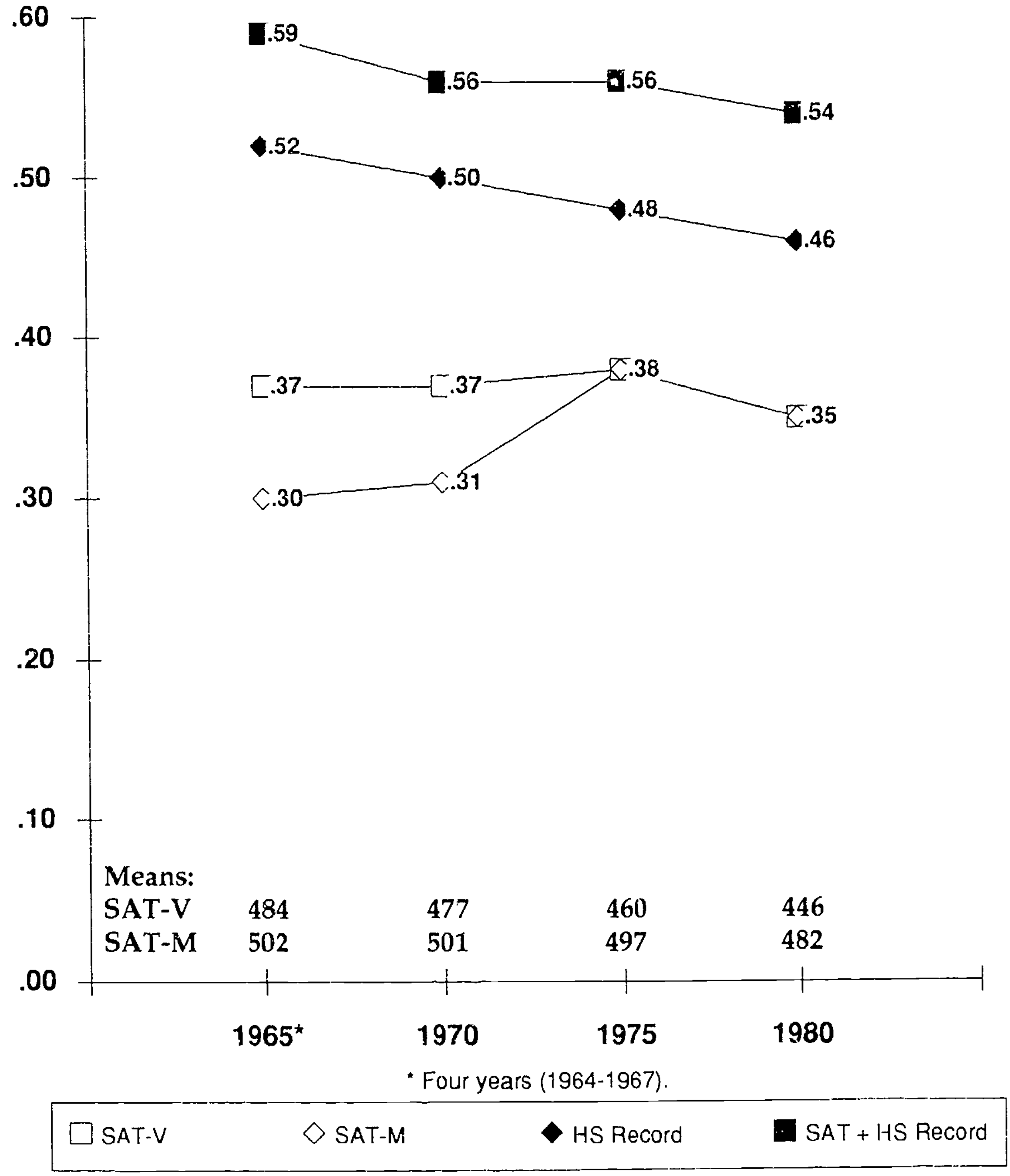
Table 2. Summary of Validity Coefficients By Gender
For College Board/VSS (1964-1981)
And USGA/CPM (1958-1983)

	SAT-V	SAT-M	V+M	HSA	HSA+SAT
CB/VSS Male Students					
Correlation			.38	.44	.51
Variance(%)			14	19	26
CB/VSS Female Students					
Correlation			.46	.49	.57
Variance(%)			21	24	32
USGA/CPM Male Students					
Correlation	.34	.38		.48	.58
Variance(%)	12	14		23	34
USGA/CPM Female Students					
Correlation	.45	.46		.54	.63
Variance(%)	20	21		29	40

Source: "The Predictive Contribution of the SAT in a Statewide System of Public Higher Education." In Measures in the College Admissions Process: A College Board Colloquium. New York: CEEB, 1986. [Original source of CEEB/VSS data: Donlon, T.F. (Ed.) The College Board Technical Handbook (1984)].

Figure 1.

Five-Year Averages in Validity Coefficients for High School Record and SAT Scores for 1964-1982 from CEEB Validity Study Service



As shown in Figure 2 the average multiple correlation for HSA+SAT with FGPA has declined from a coefficient of +.63 in 1960 to +.58 in 1985. The average correlation of HSA with FGPA has remained relatively stable, varying no more than three coefficient points in any one five-year period.

The average coefficient for SAT-V has declined from +.43 to +.32 and reflects a downward trend that recovered during the 1970s and then continued during the 1980s. A similar decline is seen in the average coefficients for SAT-M where the correlation between mathematical ability and freshman grades has changed from +.41 to +.36 for an overall decline of five coefficient points.

When SAT-V validity coefficients are analyzed by institutional level, their decline is more uneven and more noticeable. Average coefficients for four-year colleges do not show as much variation as those for two-year colleges and universities within USGA. The steepest decline, however, appears to be in USGA's three historically black institutions where the average validity coefficient drops from +.38 to +.23, a decline of fifteen points.

Changes in the validity coefficients of SAT-M suggest a pattern like those of SAT-V but with one or two sharper turns. Two-year and four-year colleges show a decline of six or seven coefficient points from 1960 to 1965, an increase of seven or eight points over the next ten years, and then a drop of five or ten points in the last decade. Universities show a steady increase over the first two decades and a decline of nine points since 1975. Historically black institutions show a drop of six points in 1980 and recovery of four points by 1985. The 1965 SAT-M Validity for two-year and four-year colleges may be attributable to increased enrollments during the 1960s.

Unlike the stability of coefficients shown for the total USGA group, the validity coefficients for HSA display one or two sharp turns when analyzed by institutional type. Four-year colleges, for example, gained six coefficient points in 1975 and lost three points over the next decade. Historically black institutions show a gain of five points from 1980 to 1985 while universities show a one point gain during the same period.

Multiple correlation coefficients for HSA+SAT show a fairly consistent decline for the four groups of institutions. The most noticeable exceptions are a four-point gain in 1975 for four-year colleges and a four point gain in 1985 for HBIs. In brief, the combination of SAT scores and high school grades as a predictor of college grades shows a decline of eight, seven, and five coefficient points for four-year colleges, two-year colleges, and universities, respectively.

Institutional Profiles

The analysis of trends by institution gives a more informative summary of the SAT's predictive validity when the SAT is used in conjunction with HSA to predict freshman GPA. The 1983 study of USGA adult learners (Fincher, 1983) suggested that college grades are more predictable when confined to a particular institution. Although the University System of Georgia has a uniform grading system, variations in academic standards and deviations from grading policies are

Figure 2.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores Averaged Across USGA Institutions (N=18)

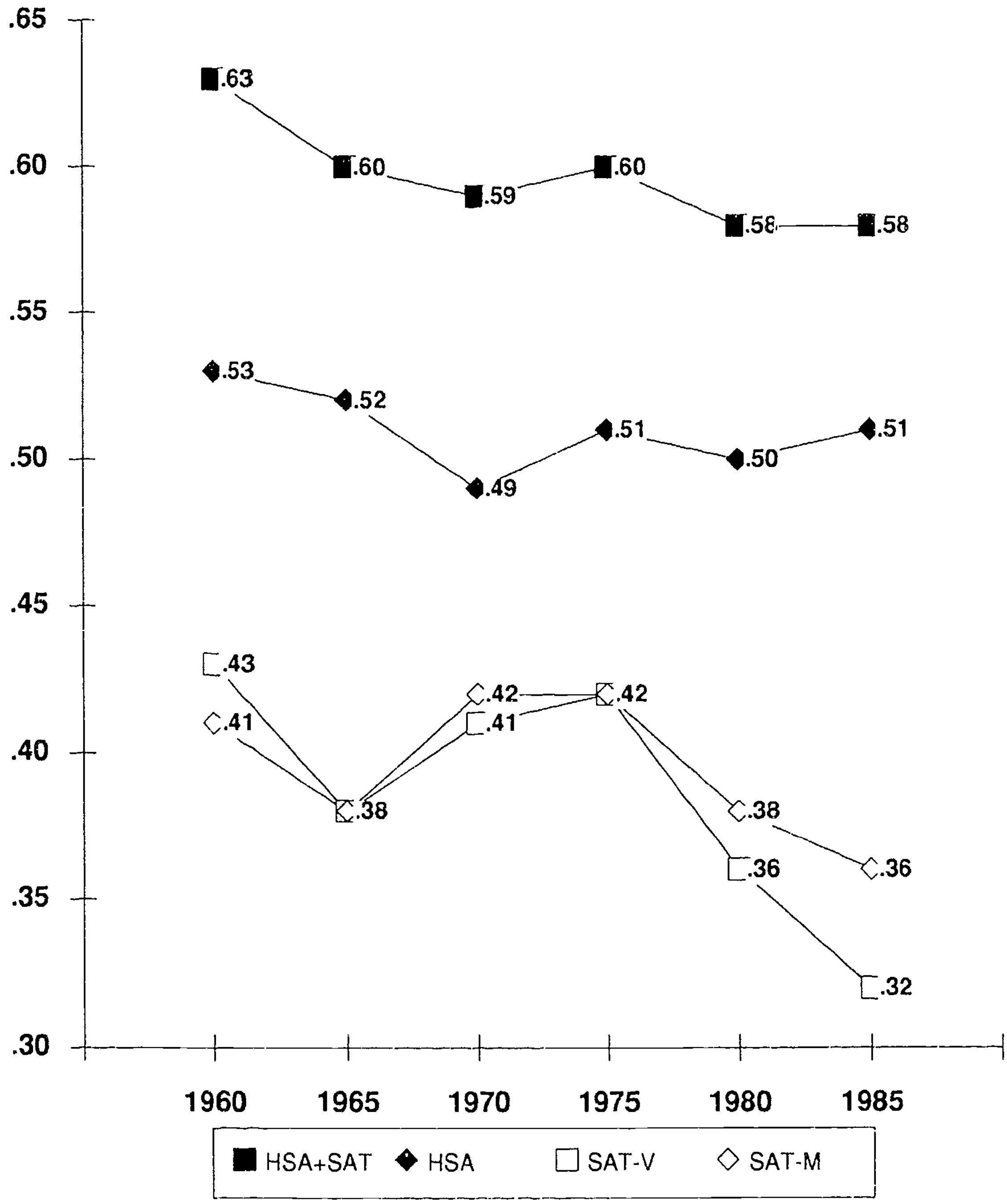


Figure 3.

Five-Year Averages in Validity Coefficients for SAT-V Scores By Institutional Level

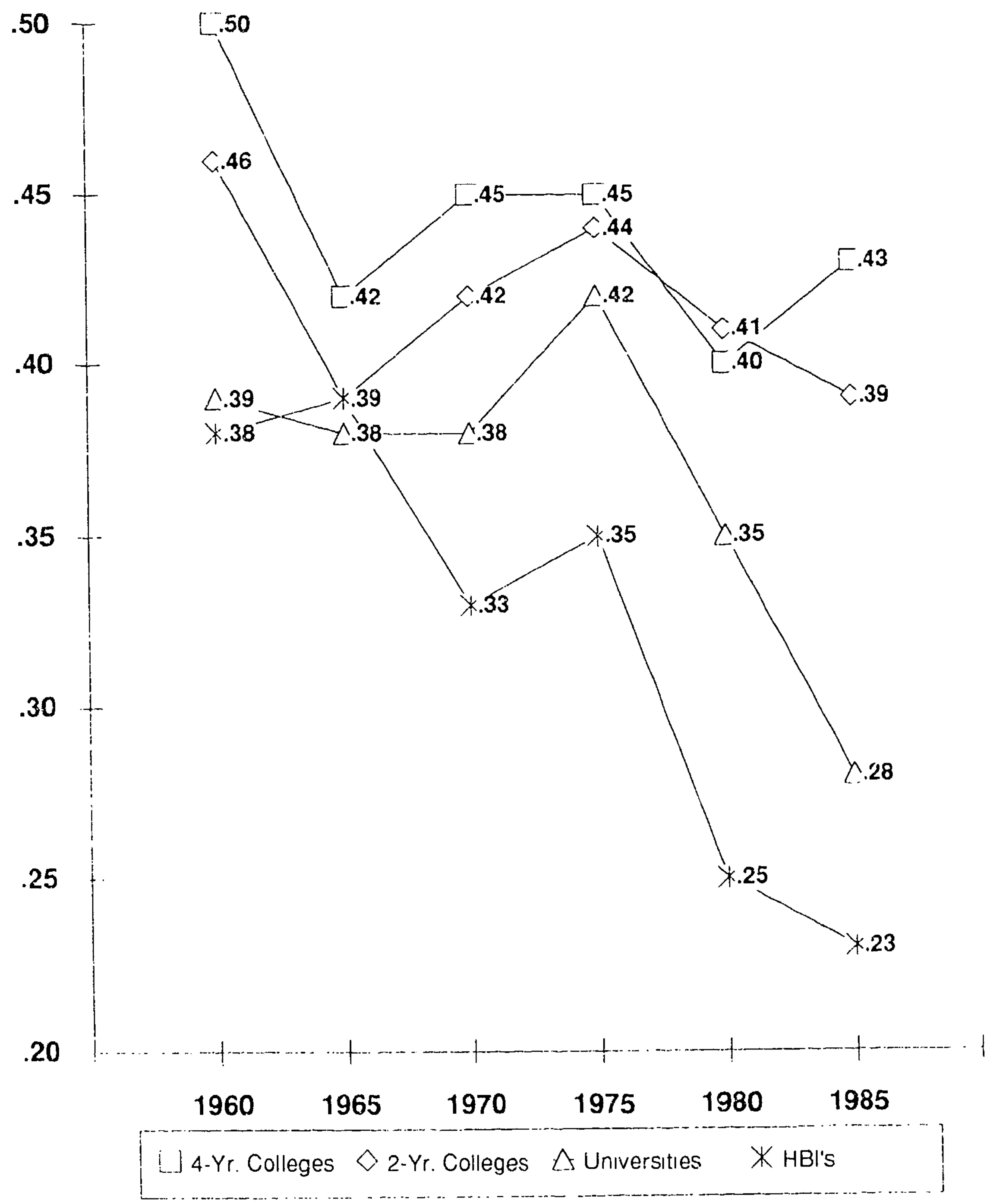


Figure 4.

Five-Year Averages in Validity Coefficients for SAT-M Scores By Institutional Level

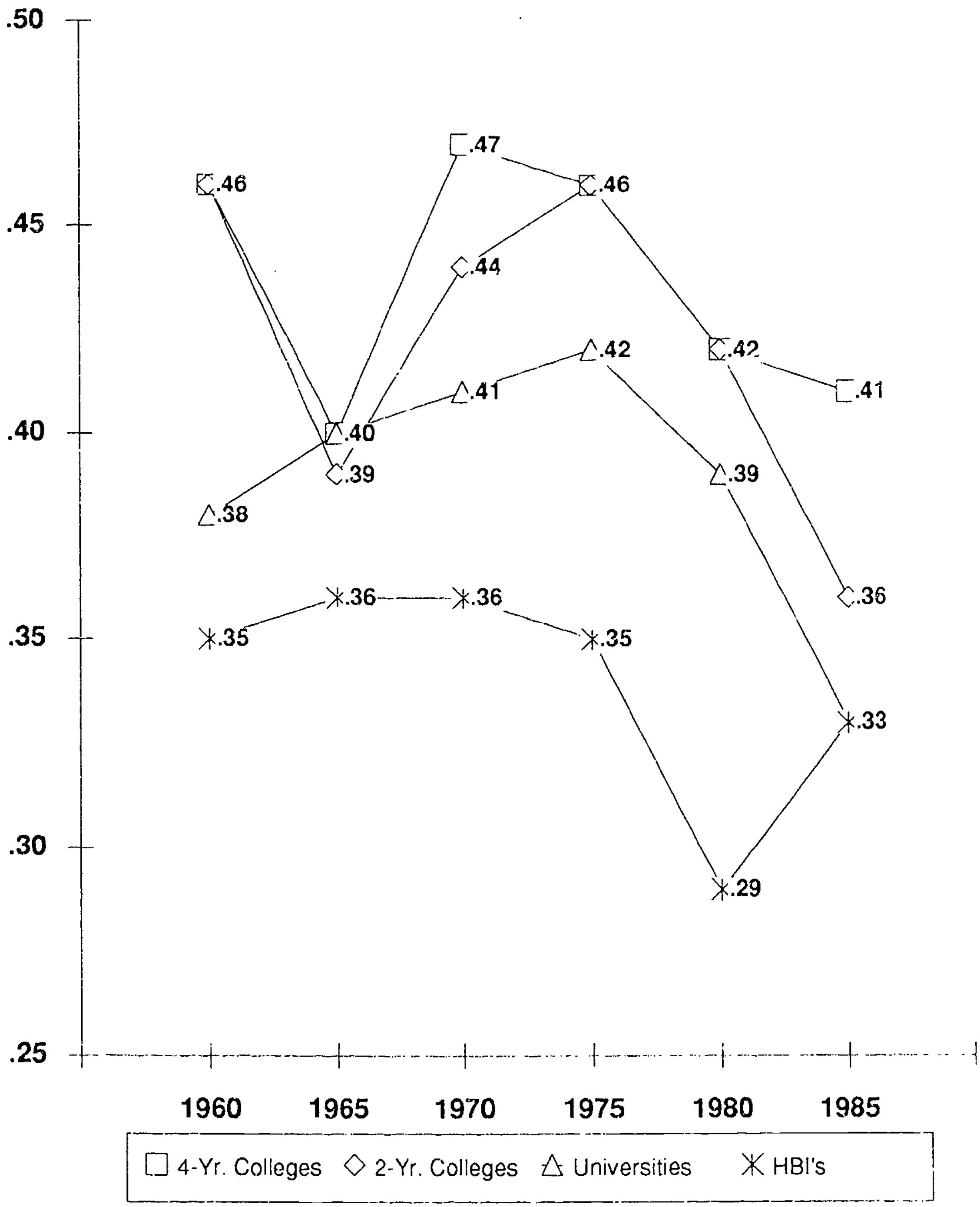


Figure 5.

Five-Year Averages in Validity Coefficients for HSA By Institutional Level

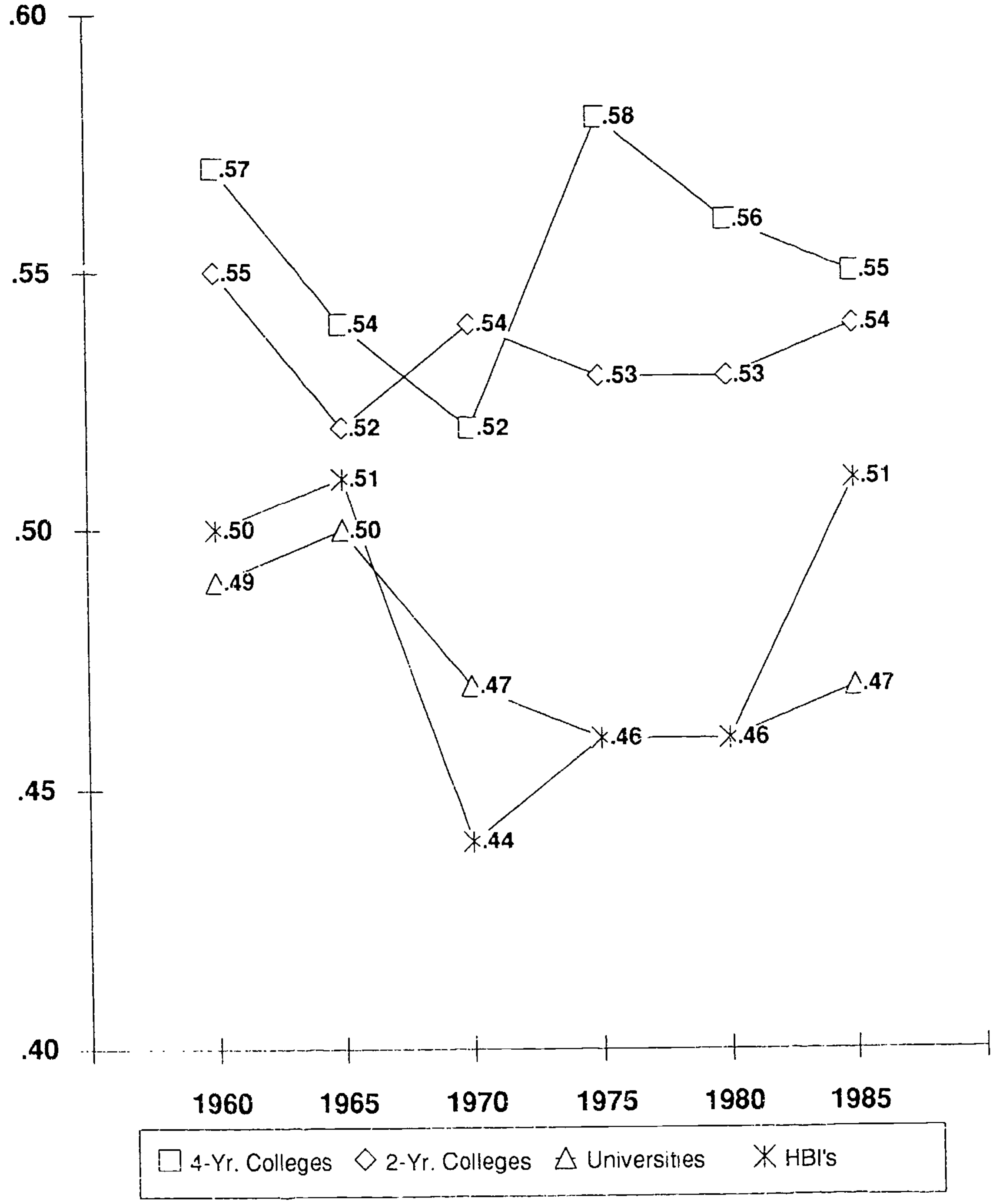
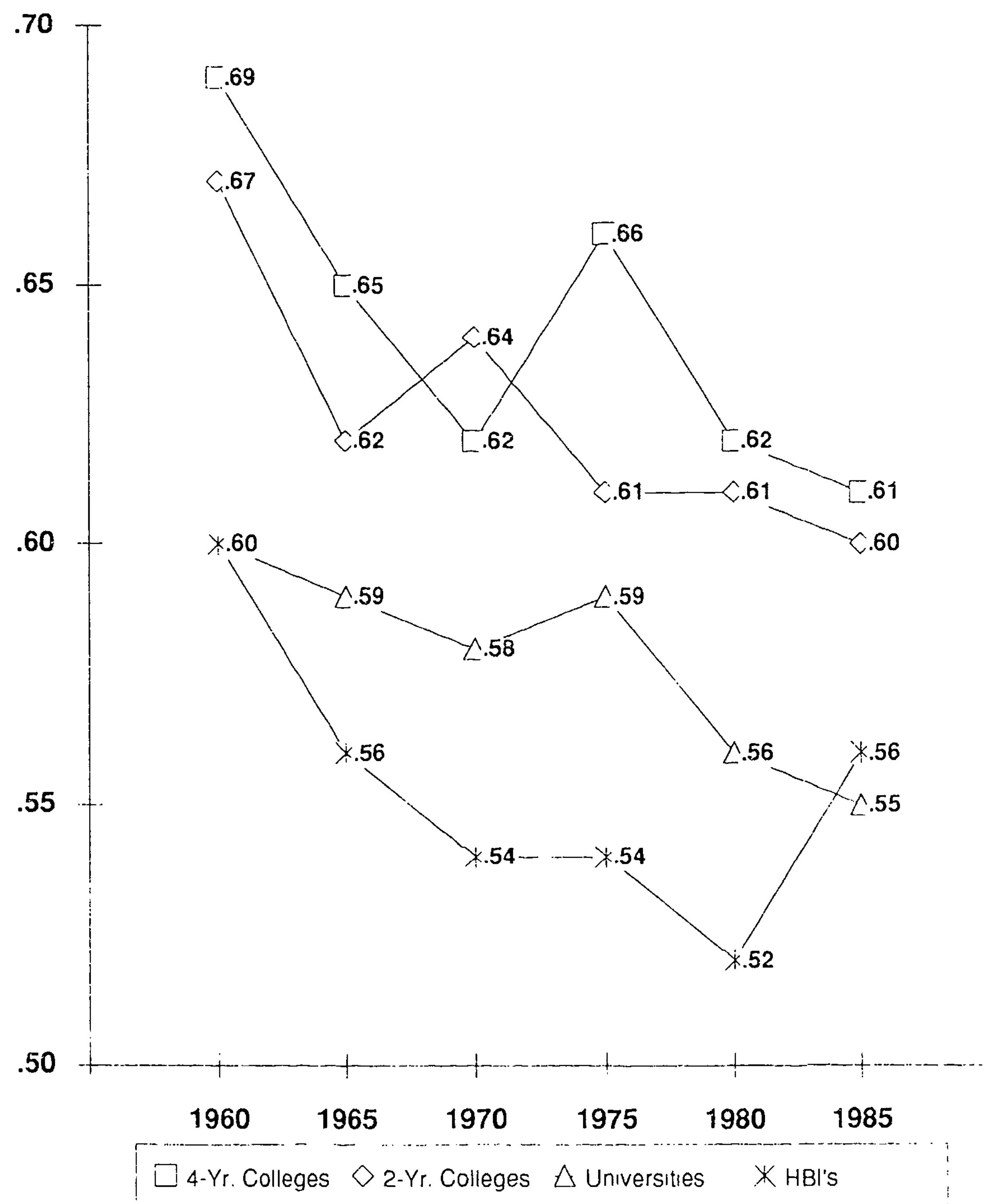


Figure 6.

Five-Year Averages in HSA + SAT Multiple Coefficients
By Institutional Level



to be expected. When SAT scores are correlated with grades that cut across institutions, there is less certainty about grading policies and practices--and less confidence in the FGPA as a measure of academic performance. Thus, the analysis of validity coefficients by institution is a way of controlling for institutional differences that might affect the correlation of predictor and criterion variables. Figures UN-1 through HBI-3 in Appendix A display institutional trends for USGA institutions.

Changes Over Ten-Year Periods

Another view of the changes in validity coefficients can be obtained from the regression slopes that have been computed for SAT-V, SAT-M, HSA, and SAT+HSA during the three ten-year periods for which data are available. The fitting of a straight line to annual validity coefficients over ten-year periods accentuates the general direction of change but disguises the unevenness of validity coefficients, when analyzed as trends.

Further information is provided by analyzing slopes and average means by gender. Data for both men and women students are available from the USGA norms booklets, and such data have been used in the analysis of ten-year trends. Institutional data vary, however, and the correlation of validity coefficients with years is meaningful only when interpreted in terms of coefficient points and when interpreted for particular institutions.

The University of Georgia will serve as an example. For male students in the 1958-1967 period, the validity coefficients for SAT-V declined 1.48 points a year while validity coefficients for SAT-M declined 1.45 points. In the 1968-1977 period coefficients for SAT-V rose 1.17 points on the average and those for SAT-M rose 1.96 points. And in the 1978-1987 period coefficients declined 0.28 points (SAT-V) and 0.50 points (SAT-M). Thus a pattern of decline, increase, and decline is suggested for three ten-year periods in which the average means of SAT-V rose from 445 to 485 and then dropped to 481--and SAT-M means showed an increase from 486 to 536 to 542.

Validity coefficients for UGA female students show a decline (1958-1967) of 1.7 points (SAT-V) and 1.0 points (SAT-M), then an increase (1968-1977) of 1.0 points (SAT-V) and 1.8 points (SAT-M), followed (1978-1987) by an increase of 0.07 points (SAT-V) and a decrease of 0.13 points (SAT-M). Thus, the pattern suggested for UGA women is decline, gain, and apparent stabilization.

The patterns suggested for HSA and SAT+HSA vary slightly from those for SAT-V and SAT-M. For UGA men the pattern for HSA and SAT+HSA of decline-and-gain diverges in 1978-1987 with a slight increase for HSA and a slight decrease for SAT+HSA. The pattern for women is more consistent: both HSA and SAT+HSA show a decrease, then an increase, and then another decrease.

In an effort to simplify detailed data, regression slopes were computed but are not reported. Without further information about public and institutional policies, shifts in academic standards, etc., interpretations based on the magnitude of change within ten-year periods are misleading. Perhaps the better informed interpretation of changes in ten-year slopes for separate institutions is the conclusion that data averaged across institutions and across years conceals the ups and downs of annual change in statistical indices such as coefficients of correlation.

Given the difficulties of interpretation, the computation of slopes, intercepts, and average predictor means has indicated that analyses over a thirty-year span of educational change are too cumbersome for the data available. It makes more sense to confine further analyses to the years following the rapid expansion of higher education in the 1960s. Undoubtedly these were disjunctive years in more ways than one, and longitudinal studies that would bridge the 1960s can be conducted only with great difficulty.

Changes in Variability

Observed variations in the predictive validity of the SAT obviously are related to the variability of student abilities, admission policies and academic standards (as those standards are reflected in grading practices). Significant changes have been observed in student abilities since the 1960s as public policies of access and equity have attracted new or different students to college campuses. Admission policies have changed substantially as entrance requirements have been modified to accommodate minority students and adult learners. And academic standards have changed in various ways as colleges have diversified course offerings to meet the educational needs of pluralistic student groups.

During the thirty years that SAT scores and high school averages have been correlated with freshman grades in the University System of Georgia, many significant changes have taken place in institutional missions and policies, degree and course requirements, instructional objectives and methods, and grading practices. All of these are related to the predictability of college grades, but none have been studied systematically during the years that SAT scores have declined nationally. There are few reasons to doubt the basic conclusions of the Advisory Panel on SAT Score Decline (1977). The SAT continues to measure the same verbal and mathematical abilities that it has been constructed and developed to measure. In similar manner there are few reasons to doubt the many changes that have taken place in statewide systems of public higher education, such as the University System of Georgia. To some unknown extent, all institutions of higher education have diversified academic programs and courses to meet the changing expectations of a more culturally pluralistic student population.

To interpret or explain the inconsistencies of predictive validity in the University System of Georgia, the variability of student abilities, secondary preparation, and academic performance at the college level must be considered. For such considerations the SAT provides much better information on verbal and mathematical abilities than we have on scholastic and/or academic achievement. Grading policies and practices often obscure student learning and accomplishments, and the teacher judgments that comprise high school averages and freshman grade-point-averages remain a mystery to most knowledgeable researchers. Fortunately, it is possible to look at the variability of high school and college grades, as they are reported in the USGA annual norms booklets.

Variations in Scores and Grades

As shown in Figure 7, the ten-year averages computed for HSA and FGPA in USGA institutions display a remarkable evenness. During 1958-1967 the average HSA for entering freshmen was 2.80; twenty years later the average HSA was 2.98, an increase of .18 grade points. College grades were even more stable, changing from an average mean of 2.27 in 1958-1967 to an average mean of 2.35 in 1978-1987.

Considering the growth in institutions from eighteen to thirty-four, in enrollments from 6,692 to 24,541 entering freshmen, in degree programs and academic courses, etc., the stability of grading patterns when computed across years and across institutions is quite remarkable. During the same thirty-year period SAT scores increased from an average of 384 to 417 (SAT-V) and from 420 to 455 (SAT-M).

Figures 8 and 9 show the shifts in average standard deviations for USGA institutions over the three ten-year periods. Fairly obvious is an increase in the average variation of freshman GPA and a decrease in the average variation of SAT-V scores. With respect to variability, only SAT-M shows an appreciable stability in its standard deviation (i.e., a twenty-year change from 111.4 to 109.1). The average standard deviation of SAT-V shows a decrease from 105.7 to 95.6, implying that the variability of verbal ability is decreasing. The standard deviation of freshman GPA shows an increase from 0.69 to 0.75 and implies that college grades are only slightly more variable in the 1980s.

Institutional patterns suggest the same kind of stability in the average performance of students. Where there is an appreciable increase in average freshman grades, it is attributable (in most cases) to changes in institutional status that produced a different mix of entering freshmen and major fields that tend to attract better prepared students. Seven two-year colleges, for example, became senior colleges serving a much broader constituency. Georgia State, as another example, became a university with greatly expanded graduate and professional offerings. Such differences in the abilities of entering students have often been noticed and were quite evident in an earlier study (Fincher, 1974) when it was "discovered" that mean SAT-V was related to institutional status (university, senior college, junior college, etc.).

Variations in Validity Coefficients

The analysis of variance in SAT-V, SAT-M, HSA, and SAT+HSA validity coefficients by decade (1970s -v- 1980s) and by gender confirms several differences that are related to validity trends. Tables 3-10 summarize the four analyses made on the twenty-one USGA institutions for which complete data (1970-1987) are available. Significant differences in SAT-V validity coefficients have been found for gender and for the decade in which they are computed. Similar differences have been confirmed for SAT-M validity coefficients but not for HSA and SAT+HSA.

Figure 7.

Ten-Year Averages of HSA and FGPA for the University System of Georgia

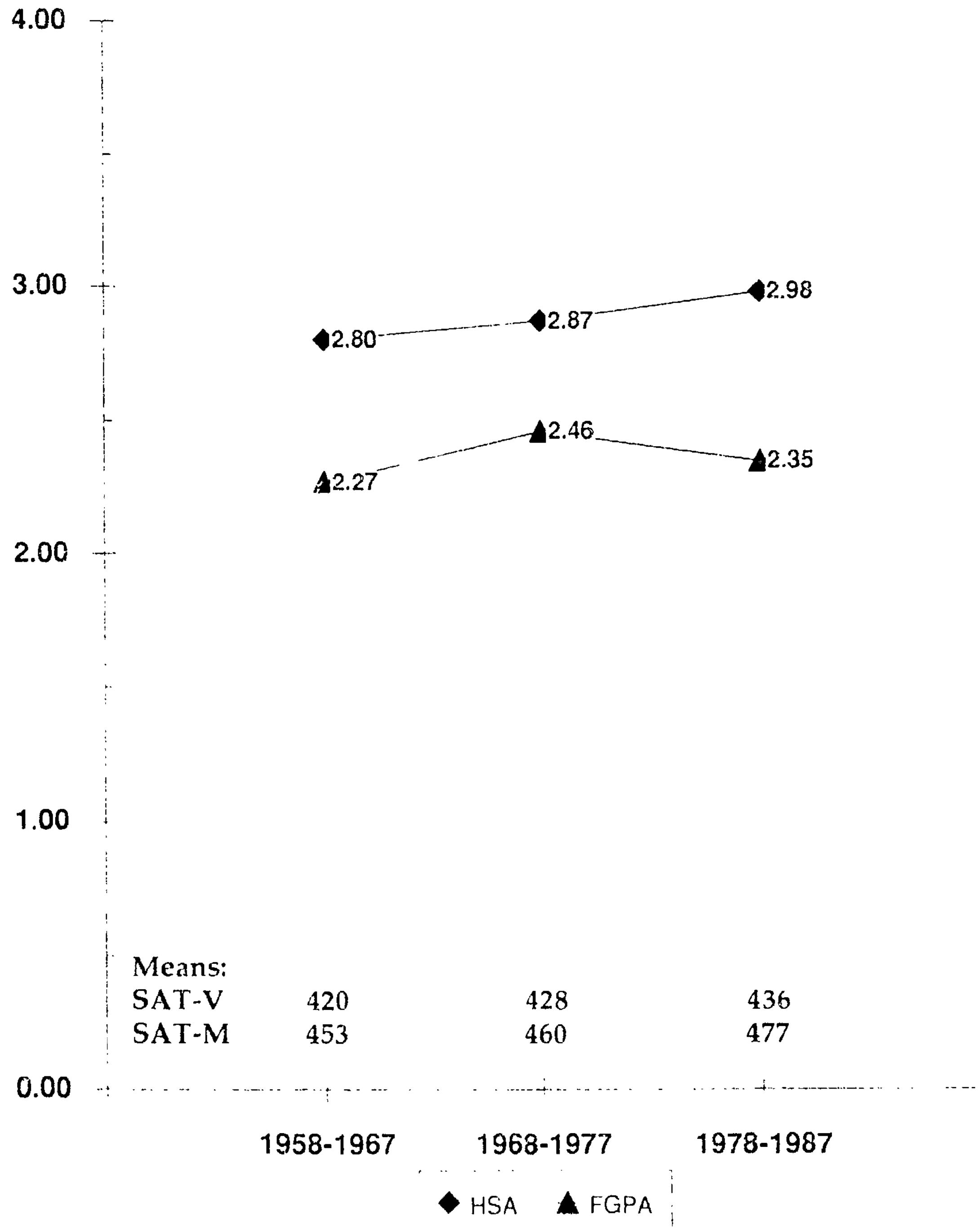


Figure 8.

Ten-Year Averages of SAT Standard Deviations for the University System of Georgia

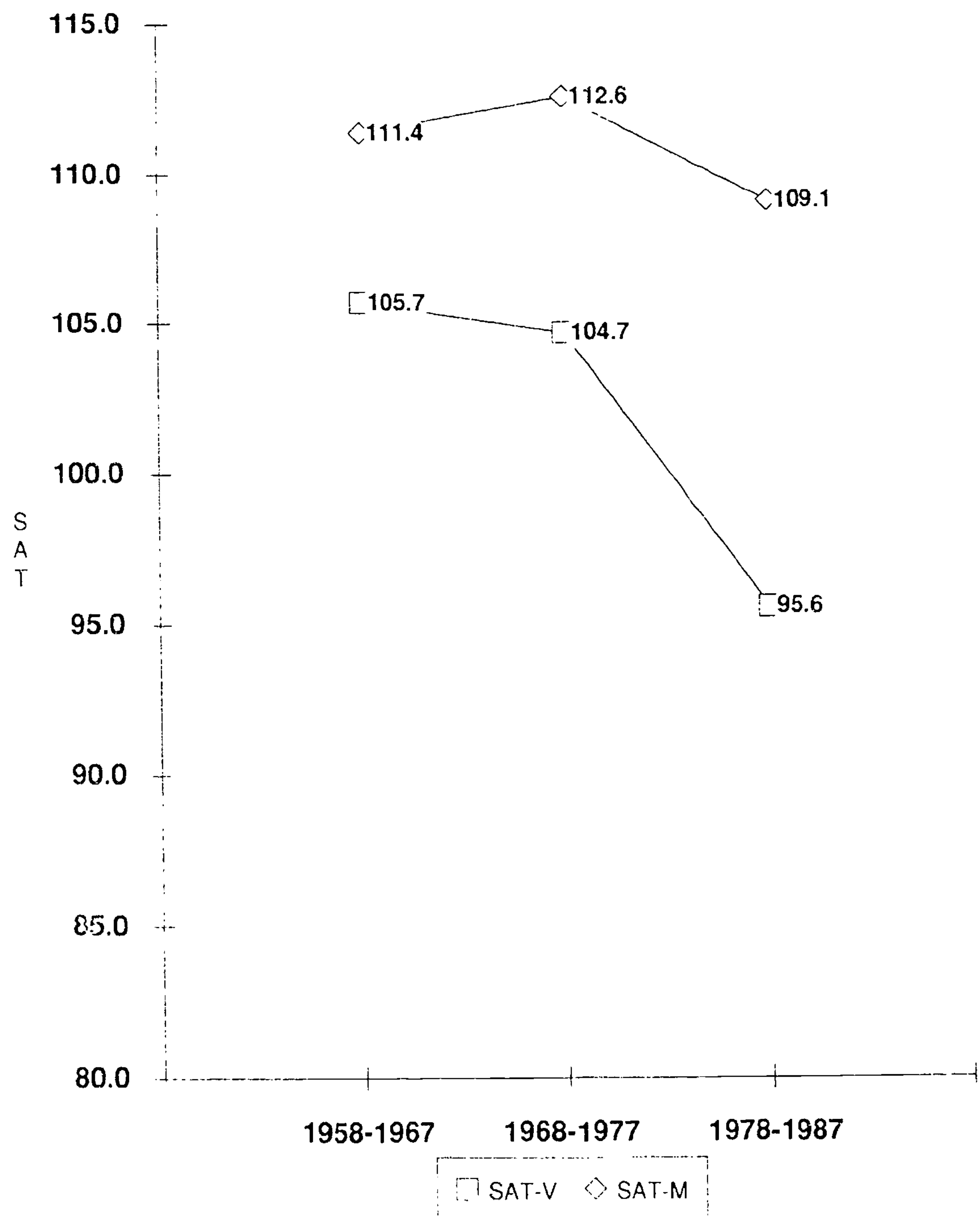
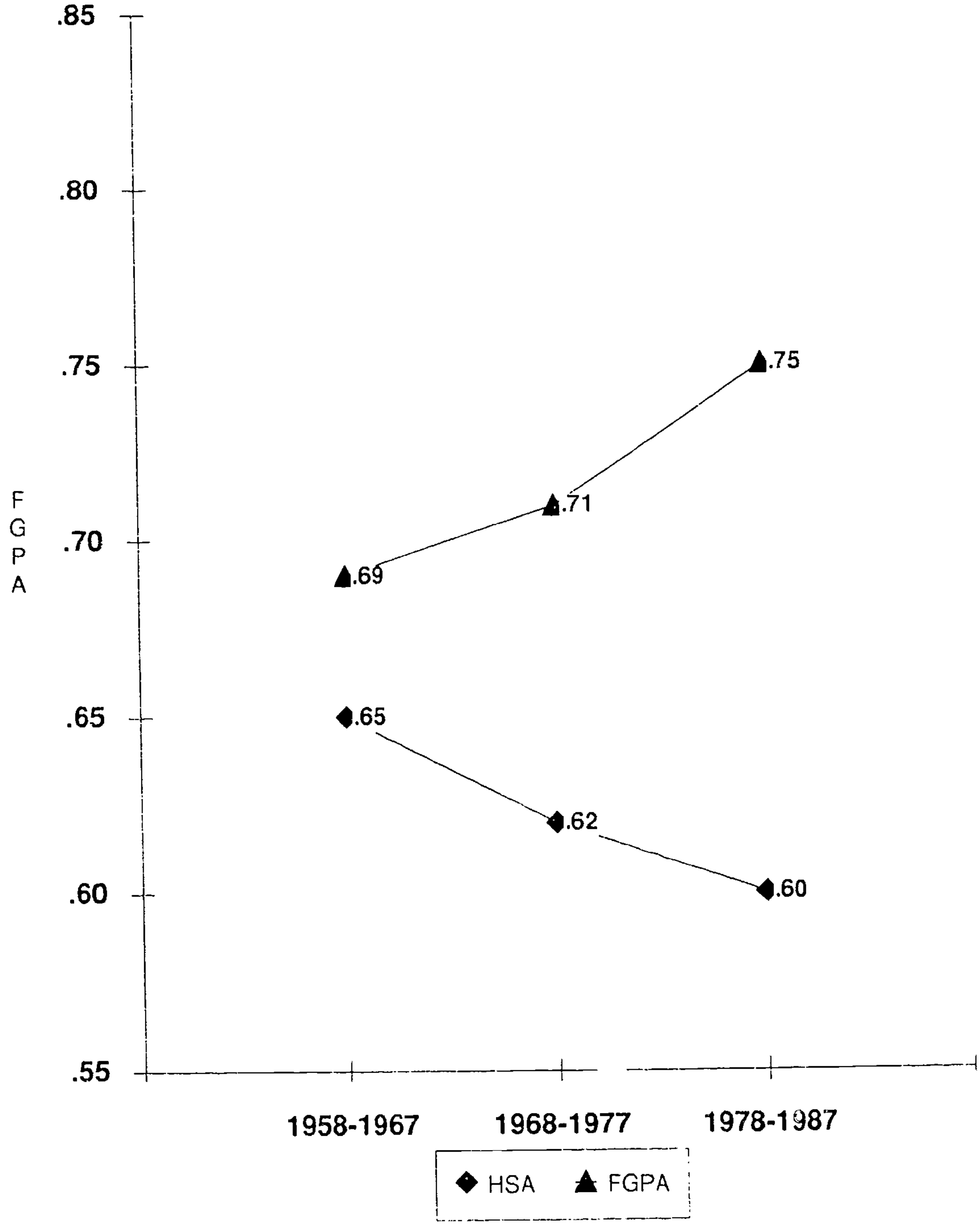


Figure 9.

Ten-Year Averages of FGPA Standard Deviations for the University System of Georgia



In terms of the additional contribution made by the SAT (when combined with HSA to predict FGPA), Table 3 indicates (for men) an additional eleven and ten coefficient points for the 1970s and 1980s, respectively. For women, the SAT contributes an additional nine and six coefficient points for the two decades. In brief, there are significant differences in the average SAT validity coefficients computed for the decades of the 1970s and the 1980s. There are also significant differences in the average coefficients computed by decade for men and women students. Such differences for HSA coefficients and SAT+HSA coefficients are not statistically significant.

In the three university-level institutions (Table 4) where the majority of USGA students enroll, decade differences in average validity coefficients are significant only on SAT-M and HSA+SAT across decades. Use of the SAT, however, contributes appreciably to the prediction of FGPA and the loss in predictive efficiency from the 1970s to the 1980s is two coefficient points.

Regional senior institutions (Table 5) show significant differences in average SAT-V and SAT-M validity coefficients for men and women for the two decades. Significant differences are also found in average HSA coefficients for men and women. For combined uses of HSA and SAT, differences in average multiple correlation coefficients are found for decades, gender, and interaction.

Four-year or community senior colleges (Table 6) also display significant differences in SAT-V and SAT-M validity coefficients. SAT scores are better predictors of FGPA for women and the decline in predictive validity from the 1970s to the 1980s is statistically significant. For the three colleges (Table 7) recently added to this group, gender differences on the SAT validity coefficients are significant, but decade differences are not. The reason for the latter finding is related, perhaps, to the change of status from junior college to senior college during the 1980s.

Gender differences in SAT-V validity coefficients are the only significant differences found in the analyses of residential junior colleges (Table 8). This finding indicates a significantly higher correlation of SAT-V scores with grade-point-averages for women.

In similar manner SAT-V validity coefficients disclose the only significant difference found in the three historically black institutions (Table 9). SAT-V scores predict grades more effectively for women than for men, and the decline in predictive efficiency from the 1970s to the 1980s is statistically significant. No significant differences have been found in the validity coefficients analyzed for SAT-M, HSA, and SAT+HSA.

For the ten junior colleges (Table 10) established at varying times since 1964, significant differences have been found in SAT-V validity coefficients (for gender and decade) and SAT-M coefficients (for gender). Predictive validity is again higher for women, and a significant decline in SAT-V predictive validity is observed in the 1970s-1980s transition. SAT-M predictive validity is higher for women, but the decade differences are not statistically significant at the .01 level of confidence.

Table 3. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for University System (N = 21 Institutions)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	29.24	.0001
Men	.35	.28	- .07	Decade	26.45	.0001
Women	.46	.39	- .07	Gender	61.26	.0001
Combined	.40	.34	- .06	D x G	.01	n.s.
SAT-M				Total	16.88	.0001
Men	.39	.34	- .05	Decade	14.73	.0002
Women	.47	.41	- .06	Gender	35.73	.001
Combined	.39	.34	- .05	D x G	.19	n.s.
HSA				Total	2.31	n.s.
Men	.47	.48	+ .01	Decade	.16	n.s.
Women	.52	.52	.00	Gender	6.65	.012
Combined	.51	.52	+ .01	D x G	.12	n.s.
HSA + SAT				Total	2.37	n.s.
Men	.58	.58	.00	Decade	1.88	n.s.
Women	.61	.58	- .03	Gender	2.62	n.s.
Combined	.59	.57	- .02	D x G	2.62	n.s.
Incremental Validity (SAT)						
Men	+ .11	+ .10				
Women	+ .09	+ .06				
Combined	+ .08	+ .05				

NOTE: Analysis based on 21 institutions for which complete data were available; "n.s." denotes p-value greater than .05.

Table 4. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Universities (UGA, GSU, TECH)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	3.52	n.s.
Men	.35	.26	- .09	Decade	6.73	n.s.
Women	.41	.33	- .08	Gender	3.72	n.s.
Combined	.36	.29	- .07	D x G	.10	n.s.
SAT-M				Total	5.64	n.s.
Men	.39	.31	- .08	Decade	10.32	.0124
Women	.44	.38	- .06	Gender	6.43	n.s.
Combined	.38	.32	- .06	D x G	.17	n.s.
HSA				Total	.37	n.s.
Men	.48	.45	- .03	Decade	.14	n.s.
Women	.45	.47	+ .02	Gender	.08	n.s.
Combined	.48	.45	+ .03	D x G	.91	n.s.
HSA + SAT				Total	3.80	n.s.
Men	.58	.54	- .04	Decade	11.38	.0097
Women	.58	.54	- .04	Gender	.00	n.s.
Combined	.59	.54	- .05	D x G	.00	n.s.
Incremental Validity (SAT)						
Men	+ .10	+ .09				
Women	+ .13	+ .07				
Combined	+ .11	+ .09				

Table 5. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Regional Senior Institutions (GSOU, VSC, WGA, GAC, NGA)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	26.56	.0001
Men	.38	.30	- .08	Decade	21.32	.0003
Women	.51	.43	- .08	Gender	58.35	.0001
Combined	.44	.36	- .08	D x G	.01	n.s.
SAT-M				Total	22.57	.0001
Men	.39	.34	- .05	Decade	14.24	.0017
Women	.51	.45	- .06	Gender	53.36	.0001
Combined	.43	.34	- .09	D x G	.10	n.s.
HSA				Total	14.46	.0001
Men	.52	.50	- .02	Decade	3.47	n.s.
Women	.57	.57	.00	Gender	39.88	.0001
Combined	.56	.54	- .02	D x G	.02	n.s.
HSA + SAT				Total	8.15	.0016
Men	.62	.61	- .01	Decade	10.93	.0045
Women	.67	.61	- .06	Gender	6.76	.0193
Combined	.64	.59	- .05	D x G	6.76	.0193
Incremental Validity (SAT)						
Men	+ .10	+ .11				
Women	+ .10	+ .04				
Combined	+ .08	+ .05				

Table 6. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Four-Year Community Colleges (ARSC, AUGC, COLC, GASW)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	23.83	.0001
Men	.41	.32	-.09	Decade	30.29	.0001
Women	.51	.42	-.09	Gender	41.16	.0001
Combined	.45	.37	-.08	D x G	.05	n.s.
SAT-M				Total	10.89	.001
Men	.44	.34	-.10	Decade	19.33	.0009
Women	.49	.42	-.07	Gender	12.69	.0039
Combined	.42	.36	-.06	D x G	.64	n.s.
HSA				Total	.39	n.s.
Men	.50	.50	.00	Decade	.10	n.s.
Women	.54	.52	-.02	Gender	1.05	n.s.
Combined	.53	.51	-.02	D x G	.03	n.s.
HSA + SAT				Total	2.23	n.s.
Men	.61	.59	-.02	Decade	4.44	n.s.
Women	.64	.59	-.05	Gender	1.12	n.s.
Combined	.62	.58	-.04	D x G	1.12	n.s.
Incremental Validity (SAT)						
Men	+.11	+.09				
Women	+.10	+.07				
Combined	+.09	+.07				

Table 7. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Four-Year Recent Colleges (CLSC, KNSC, SCT)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	7.67	.0097
Men	.30	.26	- .04	Decade	4.99	n.s.
Women	.45	.36	- .09	Gender	17.15	.0032
Combined	.33	.29	- .04	D x G	.85	n.s.
SAT-M				Total	5.94	.0196
Men	.33	.31	- .02	Decade	1.42	n.s.
Women	.42	.39	- .03	Gender	16.29	.0038
Combined	.31	.30	- .01	D x G	.12	n.s.
HSA				Total	.25	n.s.
Men	.42	.44	+ .02	Decade	.42	n.s.
Women	.44	.49	+ .05	Gender	.28	n.s.
Combined	.47	.47	.00	D x G	.06	n.s.
HSA + SAT				Total	.13	n.s.
Men	.53	.54	+ .01	Decade	.03	n.s.
Women	.56	.54	- .02	Gender	.18	n.s.
Combined	.55	.53	- .02	D x G	.18	n.s.
Incremental Validity (SAT)						
Men	+ .11	+ .10				
Women	+ .09	+ .05				
Combined	+ .08	+ .06				

Table 8. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Two-Year Residential Colleges (MDGA, SOGA, ABAC)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	12.05	.0025
Men	.38	.34	- .04	Decade	3.92	n.s.
Women	.52	.46	- .06	Gender	32.09	.0005
Combined	.44	.40	- .04	D x G	.14	n.s.
SAT-M				Total	1.82	n.s.
Men	.44	.44	.00	Decade	1.18	n.s.
Women	.51	.45	- .06	Gender	2.91	n.s.
Combined	.45	.39	- .06	D x G	1.37	n.s.
HSA				Total	.81	n.s.
Men	.50	.53	+ .03	Decade	.01	n.s.
Women	.56	.54	- .02	Gender	4.16	n.s.
Combined	.53	.55	+ .02	D x G	.95	n.s.
HSA + SAT				Total	1.72	n.s.
Men	.58	.61	+ .03	Decade	.00	n.s.
Women	.64	.61	- .03	Gender	2.57	n.s.
Combined	.61	.62	+ .01	D x G	2.57	n.s.
Incremental Validity (SAT)						
Men	+ .08	+ .08				
Women	+ .08	+ .07				
Combined	+ .08	+ .07				

Table 9. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for HBI's (ALSC, FVSC, SSC)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	14.27	.0014
Men	.27	.19	- .08	Decade	11.34	.0098
Women	.35	.31	- .04	Gender	29.87	.0006
Combined	.31	.24	- .07	D x G	1.61	n.s.
SAT-M				Total	1.57	n.s.
Men	.31	.33	+ .02	Decade	.13	n.s.
Women	.37	.36	- .01	Gender	3.83	n.s.
Combined	.32	.36	+ .04	D x G	.75	n.s.
HSA				Total	2.59	n.s.
Men	.39	.48	+ .09	Decade	3.72	n.s.
Women	.47	.49	+ .02	Gender	2.35	n.s.
Combined	.45	.50	+ .05	D x G	1.71	n.s.
HSA + SAT				Total	1.48	n.s.
Men	.51	.56	+ .05	Decade	3.21	n.s.
Women	.54	.56	+ .02	Gender	.61	n.s.
Combined	.49	.55	+ .06	D x G	.61	n.s.
Incremental Validity (SAT)						
Men	+ .12	+ .08				
Women	+ .07	+ .07				
Combined	+ .04	+ .05				

Table 10. Analysis of Validity Coefficients for SAT Scores and High School Averages by Decade and by Gender for Other Two-Year Institutions (N = 10)

	MEANS			ANOVA		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	7.79	.0004
Men	.35	.30	- .05	Decade	6.64	.0142
Women	.47	.39	- .08	Gender	16.49	.0003
Combined	.41	.33	- .08	D x G	.23	n.s.
SAT-M				Total	5.59	.003
Men	.36	.33	- .03	Decade	5.28	n.s.
Women	.45	.38	- .07	Gender	11.38	.0018
Combined	.37	.33	- .04	D x G	.11	n.s.
HSA				Total	1.95	n.s.
Men	.43	.43	.00	Decade	.10	n.s.
Women	.50	.47	- .03	Gender	5.41	n.s.
Combined	.49	.48	- .01	D x G	.35	n.s.
HSA + SAT				Total	1.05	n.s.
Men	.54	.55	+ .01	Decade	.46	n.s.
Women	.58	.55	- .03	Gender	1.35	n.s.
Combined	.57	.56	- .01	D x G	1.35	n.s.
Incremental Validity (SAT)						
Men	+ .11	+ .12				
Women	+ .08	+ .08				
Combined	+ .08	+ .08				

NOTE: Two-year colleges with data for varying years. This data not included in Table 3.

Variability and Predictive Validity

The variability of SAT-V, SAT-M, HSA, and FGPA and its respective influences on predictive validity have been analyzed by grouping USGA institutions according to their level of variability on SAT scores, high school grades, and freshman grade-point-averages. The average standard deviations of SAT-V, SAT-M, HSA, and FGPA were computed (across years) and then arrayed for classification as high-variance or low-variance institutions. Those institutions found above the median average standard deviation for USGA were placed in the high-variance category while those below the median were put in the low-variance. Institutions falling directly on the median were randomly assigned to one of the two categories. The variances of average validity coefficients were then analyzed by level of variability and by decade to determine the significance of differences among institutions in the four categories.

The results of the analysis are presented in Tables 11-15. The level of variation in SAT-V does not affect the validity coefficients of HSA and SAT+HSA but it does affect SAT-V and SAT-M validity coefficients, as researchers would logically expect. Thus, Table 11 discloses significant differences in the predictive validity of SAT scores when institutions are classified by the size of their average standard deviations in SAT-V. Also evident are significant differences in the two decades (1970s and 1980s) chosen for purposes of this study.

According to Table 12, level of variation in SAT-M affects the validity coefficients of SAT-V and SAT-M but not HSA. A significant difference, however, is found in HSA+SAT for institutions grouped by size of SAT-M standard deviation. The same institutions do <u>not</u> show a significant difference across decades for HSA+SAT.

With respect to institutions grouped by level of variation in HSA, significant differences are found for SAT-V validity coefficients by level and by decade. Significant differences are found in SAT-M coefficients for decade (and for level, if a p-value of .0215 is accepted). Significant differences in HSA validity coefficients are not found between the 1970s and the 1980s for institutions varying in HSA. The same is true for multiple correlation coefficients when HSA standard deviations are used to classify institutions; there are differences between institutions by level of variation but not between decades.

Levels of variation in FGPA display the most consistent findings. The validity coefficients of SAT-V, SAT-M, and HSA+SAT are affected by variance in freshman GPA but the validity coefficients of HSA are not. The observed pattern of significant differences thus supports an interpretation that changes in validity coefficients are attributable, in part, to the variability of college grades. HSA, as an index of secondary preparation, apparently accommodates the changes that are taking place in student abilities. FGPA, as an index of academic achievement, is also "accommodating," but restriction of range in college grades is nonetheless evident and nonetheless influential in the predictive validity of the SAT.

Table 11. Analysis of Variance in Validity Coefficients By Decade By Level of Variation for SAT-V For the University System of Georgia

	MEANS			ANALYSIS		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	6.43	.0012
Hi-Var	.42	.36	-.06	Variance	7.24	.0105
Lo-Var	.37	.37	.00	Decade	11.91	.0014
DIFF	-.05	+.01		V x D	.16	n.s.
SAT-M				Total	4.41	.0093
Hi-Var	.40	.37	-.03	Variance	5.18	.0286
Lo-Var	.38	.31	-.07	Decade	7.32	.0101
DIFF	-.02	-.06		V x D	.73	n.s.
HSA				Total	.46	n.s.
Hi-Var	.52	.52	.00	Variance	1.38	n.s.
Lo-Var	.50	.49	-.01	Decade	.01	n.s.
DIFF	-.02	-.03		V x D	.01	n.s.
HSA + SAT				Total	2.38	n.s.
Hi-Var	.61	.59	-.02	Variance	5.69	.0221
Lo-Var	.57	.55	-.02	Decade	1.44	n.s.
DIFF	-.04	-.04		V x D	.02	n.s.
Incremental Validity (SAT)						
Hi-Var	+.09	+.07	-.02			
Lo-Var	+.07	+.06	-.01			
DIFF	+.02	+.01				

Table 12. Analysis of Variance in Validity Coefficients By Decade By Level of Variation for SAT-M For the University System of Georgia

	MEANS			ANALYSIS		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	15.75	.0001
Hi-Var	.44	.37	- .07	Variance	29.50	.0001
Lo-Var	.35	.28	- .07	Decade	17.71	.0002
DIFF	- .09	- .09		V x D	0.00	n.s.
SAT-M				Total	7.95	.0003
Hi-Var	.42	.36	- .06	Variance	13.31	.0008
Lo-Var	.35	.32	- .03	Decade	9.92	.0032
DIFF	- .07	- .04		V x D	.63	n.s.
HSA				Total	.88	n.s.
Hi-Var	.52	.51	- .01	Variance	2.46	n.s.
Lo-Var	.49	.49	.00	Decade	.01	n.s.
DIFF	- .03	- .02		V x D	.16	n.s.
HSA + SAT				Total	2.66	n.s.
Hi-Var	.61	.59	- .02	Variance	6.38	.0158
Lo-Var	.57	.55	- .02	Decade	1.46	n.s.
DIFF	- .04	- .04		V x D	.15	n.s.
Incremental Validity (SAT)						
Hi-Var	+ .09	+ .08	- .01			
Lo-Var	+ .08	+ .06	- .02			
DIFF	+ .01	+ .02				

Table 13. Analysis of Variance in Validity Coefficients By Decade By Level of Variation for HSA For the University System of Georgia

	MEANS 1970s	1980s	DIFF	ANALYSIS	F-ratios	p-levels
SAT-V				Total	8.74	.0002
Hi-Var	.43	.36	- .07	Variance	12.77	.0010
Lo-Var	.36	.30	- .06	Decade	13.34	.0008
DIFF	- .07	- .06		V x D	.09	n.s.
SAT-M				Total	4.40	.0094
Hi-Var	.40	.37	- .03	Variance	5.75	.0215
Lo-Var	.37	.31	- .06	Decade	7.32	.0102
DIFF	- .03	- .06		V x D	.13	n.s.
HSA				Total	3.71	.0195
Hi-Var	.53	.54	.01	Variance	10.94	.0021
Lo-Var	.48	.47	- .01	Decade	.01	n.s.
DIFF	- .05	- .07		V x D	.19	n.s.
HSA + SAT				Total	5.33	.0037
Hi-Var	.61	.60	- .01	Variance	14.14	.0006
Lo-Var	.56	.54	- .02	Decade	1.72	n.s.
DIFF	- .05	- .06		V x D	.12	n.s.
Incremental Validity (SAT)						
Hi-Var	+ .08	+ .06	- .02			
Lo-Var	+ .08	+ .07	- .01			
DIFF	.00	- .01				

Table 14. Analysis of Variance in Validity Coefficients By Decade By Level of Variation for FGPA For the University System of Georgia

	MEANS			ANALYSIS		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	9.33	.0001
Hi-Var	.44	.34	- .10	Variance	11.29	.0018
Lo-Var	.35	.31	- .04	Decade	13.71	.0007
DIFF	- .09	- .03		V x D	2.98	n.s.
SAT-M				Total	8.20	.0002
Hi-Var	.43	.34	- .09	Variance	5.96	.0194
Lo-Var	.35	.34	- .01	Decade	10.03	.0030
DIFF	- .08	.00		V x D	8.61	.0056
HSA				Total	4.14	.0124
Hi-Var	.54	.49	- .05	Variance	2.28	n.s.
Lo-Var	.47	.52	+ .05	Decade	.01	n.s.
DIFF	- .07	+ .03		V x D	10.14	.0029
HSA + SAT				Total	4.81	.0062
Hi-Var	.63	.56	- .07	Variance	4.53	n.s.
Lo-Var	.55	.57	+ .02	Decade	1.67	n.s.
DIFF	- .08	+ .01		V x D	8.23	n.s.
Incremental Validity (SAT)						
Hi-Var	+ .09	+ .07	- .02			
Lo-Var	+ .08	+ .05	- .03			
DIFF	+ .01	+ .02				

Table 15. Analysis of Variance in Validity Coefficients By Decade By Level of Variation for Select Group FGPA For the University System of Georgia

	MEANS			ANALYSIS		
	1970s	1980s	DIFF		F-ratios	p-levels
SAT-V				Total	9.28	.0001
Hi-Var	.44	.36	- .08	Variance	12.57	.0012
Lo-Var	.37	.31	- .06	Decade	14.66	.0006
DIFF	- .07	- .05		V x D	.63	n.s.
SAT-M				Total	7.12	.0009
Hi-Var	.43	.35	- .08	Variance	6.11	.0190
Lo-Var	.36	.33	- .03	Decade	13.01	.0010
DIFF	- .07	- .02		V x D	2.23	n.s.
HSA				Total	4.52	.0094
Hi-Var	.54	.53	- .01	Variance	13.04	.0010
Lo-Var	.48	.48	.00	Decade	.29	n.s.
DIFF	- .06	- .05		V x D	.23	n.s.
HSA + SAT				Total	4.69	.0080
Hi-Var	.63	.59	-.04	Variance	8.32	.0070
Lo-Var	.58	.55	- .03	Decade	5.22	.0291
DIFF	- .05	- .04		V x D	.52	n.s.
Incremental Validity (SAT)						
Hi-Var	+ .09	+ .06	- .03			
Lo-Var	+ .10	+ .07	- .03			
DIFF	- .01	- .01				

Profiles of Cooperating Institutions

To gain a better perspective on SAT validity trends, an inquiry was made by mail or phone to over 200 other colleges and universities. The objectives of the study were identified in a covering letter and a report form was enclosed for data and information about validity trends that were available from independent institutional studies of the SAT. Institutions were selected on the basis of personal knowledge and contacts (and on the likelihood that the selected institutions would have data that could be used in the study).

Varying responses were received from fifty-three institutions, either by mail, phone, or BITNET. Longitudinal data were obtained from fourteen institutions that supplied unusually good information for inclusion in the study. Other institutions supplied useful information but with data for two years or less. Others provided information that reflects institutional or campus conditions related to uses of the SAT in predicting academic performance. For examples, two colleagues confided that their institutions conducted validity studies for internal use only and could not reveal their study findings. Other respondents indicated that state laws and governing board policies were the basis of institutional policies that precluded their participation in the study.

Figures CI-1 through CI-14B in Appendix B give the validity trends for fourteen institutions supplying longitudinal data. Although there are variations in institutional use of high school rank, high school average, and SAT scores weighted differentially or combined arithmetically, the patterns revealed by institutional profiles are often similar. Whatever the rise and fall of one predictor's validity coefficients, others appear to follow a parallel course. Generally speaking, declines in predictive efficiency are not pronounced unless they are viewed over an extended period of time. And on all occasions where increases in predictive validity are observed, they are best regarded as modest.

Other impressions suggest that cooperative institutions have neither reported a serious decline in the predictive validity of SAT scores nor a remarkable constancy in the efficiency of high school grades. Both SAT and HSA suggest annual variations in predictive efficiency and/or effectiveness. Implied in some of the profiles is the likelihood that entering students are not as homogeneous as they were once thought to be.

The summaries of the profiles and summaries developed for cooperating institutions have provided additional impressions about the uses of the SAT in the responding institutions and additional insight into the changes that are taking place in validity coefficients.

At least one generalization is undisturbed by the profiles: When combined, high school grades and SAT scores produce a more accurate prediction of freshman academic performance than either can alone. And in an era of remarkable computational capabilities, there are no known reasons why SAT-V, SAT-M, and HSA should not be differentially weighted routinely in predicting freshman GPA.

Conclusions and Implications

This study has produced many graphs and tables that depict variations in validity coefficients over a varying number of years. From these graphs and tables, we may conclude: (1) an appreciable decrease in validity coefficients can be observed for SAT-V and SAT-M during the past six to seven years, (2) the decrease in SAT validity coefficients is accompanied by similar variations for HSA coefficients and for multiple correlation coefficients involving combinations of HSA and SAT, (3) the same general trends are observed when annual variations are smoothed by computing five-year averages and ten-year averages, and (4) analyses have confirmed that the general trends are affected by the variability of student abilities, as measured by the SAT, and academic performance, as reflected in high school and college grades.

Efforts to analyze the observed variations in validity coefficients lead to the following conclusions: (5) general trends are best depicted in the computation and plotting of five-year averages for SAT-V, SAT-M, HSA, and SAT+HSA validity coefficients; (6) efforts to analyze these general trends have been facilitated by the computation of ten-year averages in validity coefficients; and (7) confinement of the analyses to variations in ten-year averages for the 1970s and the 1980s provide the most meaningful information about possible determinants of the observed changes in validity coefficients.

All such findings of this study are subject to interpretation. Decreases in the magnitude of validity coefficients do not mean a decline in the meaning and significance of the SAT as a measure of verbal and mathematical abilities. It is distinctly possible that the construct and content validity of the SAT remains intact, and the observed changes in validity coefficients are a function of other changes taking place in secondary and higher education. Annual variations do imply, however, a certain amount of instability in correlation coefficients that could be attributable to the individual differences of successive classes of entering freshmen. Annual variations, therefore, are a source of error in the prediction of freshman GPA for particular classes.

To interpret the general trend of validity coefficients, variations in four predictor variables (SAT-V, SAT-M, HSA, and SAT+HSA) and a single criterion variable (FGPA) have been extensively analyzed for units of the University System of Georgia, a statewide system of public higher education that has required the SAT for admission over a thirty-year period. Intensive efforts have been directed to the USGA data because predictive uses of the SAT are dependent upon the comparability of normative data, regression coefficients, and institutional policies. In using a regression equation based on previous classes, admissions policies must assume that applicants under consideration are comparable to the enrolled students on whom regression equations are based. Also necessary is the assumption that changes in high school and college grading practices have not changed sufficiently to alter the predictability of future student performance. Given the changes in validity coefficients that are evidenced by this study, the comparability (from year to year) of high school and college grades is a matter inviting more intensive study.

The collection and analyses of data for this study have underscored the frequent observation that changes in validity coefficients are affected by changes in the variability of student abilities and performance. In turn, changes in test scores and grades reflect larger, more significant changes in student and school characteristics. The apparent "stability" of HSA, as indicated in many graphs, suggests that high school standards have "floated" with the mix of students graduating from high school in the 1980s. To a certain extent, the same is true of academic standards at the college level. The FGPA, as a criterion of academic performance, accommodates a different mix of students taking freshman coursework in public institutions committed to increased access and equity. In other words, grading standards and practices are affected subtly (if not directly) by the mix of students now attending college.

This study does not (and indeed cannot) verify either stability or decline in the validity of SAT scores per se. There are implications within these data, however, that the SAT continues to measure what it was intended to measure and what it has measured over a sixty-year period--namely the verbal and mathematical abilities of individual students entering college. What the SAT predicts (i.e., college grades) may have undergone significant change.

Changing Conditions

This study has been conducted for a period of years in which numerous changes in public and institutional policies have taken place. Within the USGA, as only the most obvious examples, there have been changes which could and which in all probability have directly affected the validity coefficients as they are now computed and reported. Despite a uniform grading scale (4.0 = A) and a core curriculum that should assure (reasonably) comparable course content in the freshman year, changes in other systemwide and institutional policies may exercise greater influence on grading practices and/or academic standards. These changes may be identified briefly as:

1. The establishment of developmental studies programs within all units of the University System (with the exceptions of the Medical College and Georgia Tech). University System and institutional policies specify that developmental studies coursework is given institutional credit only and does not involve degree credit. This means that many students are not included in the computation of the validity coefficients which have been analyzed in this particular study. These policies and practices also restrict the range of SAT scores by excluding those students who have scored below 330 on either SAT-V or SAT-M. These policies and practices thus restrict the range of freshman grades by excluding all grades which are earned in developmental courses. Restriction of range, however, is preferable to the distortion of the FGPA that would result from including A's and B's earned in developmental English and mathematics with A's and B's earned in degree-credit English and mathematics courses.

2. Further restriction of range is imposed upon the normative data by the requirement that students must attempt 25 credit hours before their grade-point-averages are used in the computation of correlation coefficients and the derivation of regression equations. This effect on variability is evidenced by decreases of: (a) five to seven points in the standard deviations of SAT-V and SAT-M, (b) five or six points in the standard deviations of HSA for freshmen attempting 25 hours of coursework, and (c) .15 to .20 points in the standard deviations of freshman GPA. It is the latter decrease in standard deviations that is most impressive. This decrease in variability of freshman grades implies that restriction of range in the criterion variable may have the most telling effect on the observed reductions in validity coefficients.

3. Other changes in University System and institutional policies that might affect the magnitude of validity coefficients are such matters as the admission policies determining the entry of adult learners. As adult learners have entered units of the USGA there have been inconsistencies of both policy and application. For example, adult learners were first required to present SAT scores for admission. Then, a contingency policy was established whereby they could enter units of the University System without taking the SAT. The 1983 adult learner study indicates that the SAT does predict freshman grades for adult learners but at a lower level of predictive efficiency. Indeed, one important finding of the adult learner study was a significant correlation of SAT scores and freshman grades even when there was no HSA by which to judge the previous achievement of students entering units of the University System. All such matters aside, the inclusion of (non-traditional) adult learners and women students is a contributing factor in recent variations of validity coefficients. As more part-time, older, and non-degree-seeking students enter units of the University System and other institutions of higher education, variations in the predictability of academic performance in the first year of college are observed.

Research Perspectives

Interpretations of validity trends, as reported in this study, are also dependent upon perspective. The unit of analysis for this study has been the correlation coefficients of SAT-V, SAT-M, HSA, and SAT+HSA with freshman GPA. At least three "interpretative models" could be applied in analyzing such data: a predictive efficiency model in which attention is directed to the reduction of predictive error; a variance partitioning model in which variations in FGPA are explained by related variations in SAT-V, SAT-M, and HSA; and an analysis of effects model in which changes in SAT-V, SAT-M, and HSA are analyzed in terms of their effect on academic performance, as reflected in FGPA.

To comprehend the changes that are taking place in SAT validity coefficients and thereby in the predictive contributions of SAT, validity coefficients are believed to be the proper unit of analyses. Where indications of incremental validity seems appropriate, these are reported as increases (or decreases) in correlation coefficient points and not as the percent of additional variance explained by the inclusion of SAT scores in regression equations.

Previous experience with the SAT has shown that first order coefficients can be averaged without distorting the meaning of inter-relationship among SAT scores, high school grades and academic achievement. In other words, conversion to Fisher Z-coefficients prior to averaging correlation coefficients has not significantly altered the mean value derived by computing simple arithmetic averages of the coefficient. One reason for this is the narrow range in which most validity coefficients fall. For example, most of the validity coefficients analyzed in this study fall between +.15 and +.55, a range of forty coefficient points.

In the interpretation of reported findings, many of the changes can be regarded as small or as large, depending upon the interpreter's preferences. Critics of the SAT will undoubtedly find further justification of their preferred interpretation. They may even find justification for their stance that the SAT ought to be discontinued or substantively modified. To the contrary, there is nothing in this particular study to indicate that the SAT ought to be discontinued or that it ought to be substantially modified. If the SAT does not now predict as well as it has in the past, that minor fact is most likely due to the fact that high schools do not grade and teach students with the same standards they have applied in the past, and it is almost certain that colleges do not teach and grade in exactly the same way as they did 25 years ago. These generalizations are confined, of course, to the first year of postsecondary education where a great deal of continuing selection takes place.

An important conclusion to be drawn from this particular study is that valuable information is still provided by SAT scores, their validity coefficients, and the incremental contribution that scores make to the prediction of college grades. Another point to be made with respect to the predictive validity of the SAT is the dependence of predictive validity upon the metric properties of college grades. Freshman GPA's have some remarkable characteristics and their effectiveness as a measure of achievement is notable. Despite recently expressed concerns over grade inflation, the scalar properties of grades have not been sufficiently studied during the years of test score decline and they are not being systematically studied in 1989. Numerous studies of the SAT support the thesis that cumulative FGPA's provide a more reliable and predictable measure of student learning than first-quarter grades. Indeed, the 1983 adult learner study indicated that first-quarter grades are poor indicators of how well adult learners might do in college coursework. A quarter or two of adjustment evidently is needed and judgments of student learning based on first-quarter coursework only are premature. This interpretation is also implicit in the requirement of 25 hours credit for inclusion in the USGA norms booklets.

Another perspective--or professional opinion--is to the effect that magnitude of validity coefficients is neither the sole nor the best criterion by which to judge the effective use of the SAT in making academic decisions. There are no universal criteria by which magnitude can be judged. Validity coefficients are often modest but statistically significant and educationally relevant. To interpet adequately the validity coefficients included in this study, it is necessary to look at the ways in which SAT scores are used in admission, advising, placement, counseling, and other academic decisions and choices. The 1986 study of SAT uses in the University System indicates that the SAT has many uses which often go unnoticed by critics of the SAT. The informal responses of colleagues in cooperating institutions suggest, in many ways, that predictive validity or efficiency is not a matter of concern on the campuses of many major institutions. The SAT is still required for admission purposes and occasional studies of its effectiveness (as related to grades) may be conducted, but for the most part, SAT scores are considered as supplementary information in individual academic decisions. At least two respondents stated that "no studies of the SAT have been made since the sixties." Implied in both statements is a sense of futility in predicting grades that no longer reflect (accurately) the academic standards of the institution. Also implied by other informal responses is an increasing use of SAT scores as absolute (and independent) measures of academic ability that may or may not be related to other indices or judgments of academic performance. In brief, many institutions are not interested in the SAT's effectiveness in predicting grades.

Given the variations in validity coefficients that occur from year to year, there is no cause given by this study for alarm in the decline which can be observed in the past seven to eight years. There are good reasons to believe that the HSA does not show as great a decline because high school standards flow with the ability of students in high school. Colleges, on the other hand, seem to be accommodating individual differences in verbal and mathematical and other academic abilities by such actions as developmental studies, options for no-pass or non-punitive grading, and other forms of non-credit instruction that supplement regular academic coursework. In brief it is the educational relevance of SAT scores and their statistical significance that should be given priority over the magnitude of their validity coefficients with respect to HSA.

High school averages continue to be the best single predictor of freshman grades and they do so for many reasons. Not the least of these reasons is the commonality of content in coursework and the similarity of judgments that are made by school teachers and college instructors. In other words, whatever the principal or major components of high school courses and college freshman courses, they tend to revolve around English and social studies much more than they revolve around mathematics and technical, specialized, or advanced courses. This, no doubt, is one reason why SAT-V tends to correlate higher with FGPA than SAT-M does. The similarities of teacher judgment provide an even stronger argument. Given the challenge of teaching poorly prepared students at both the high school and the college level, many teachers use grades as a motivational strategy and will grade effort when they cannot grade achievement. In many cases this is an educationally sound practice and as it should be. The fact does account, however, for the fact that SAT scores are measuring other student characteristics and thus should not be condemned for not measuring everything.

Academic Policies and Standards

No variations in validity coefficients have been presented in this study that cannot be explained as readily by changes in academic policies and standards as changes in the predictive efficiency of the SAT. When complex relationships vary, all sides of those relationships must be questioned.

Perhaps the most important conclusion to be drawn from this study is the widespread need for a more comprehensive (quite different) study of academic policies and standards, as policies and standards relate to access and equity and to the use of standardized measures of academic ability and achievement. Changes in institutional and academic policy--and not the technical merits of standardized tests--are in need of further study. The benefits of such research are suggested by the intensity of past debates over differential validity, the rapidity with which intense debate faded, and the eventual emergence of a concern with generalized validity. Differential validity (and test bias) was not a problem that could be resolved on technical merits alone (Fincher, 1975).

In much the same manner, the predictive uses of SAT scores cannot be debated on their technical merits. All uses of the SAT are a matter of public and institutional policies that require concerted and systematic examination. Many changes in institutional and academic policy have gone unnoticed since the early 1970s. Many observers of higher education are convinced that academic standards have eroded in the face of an active, well intended concern with educationally disadvantaged students.

As a national concern with the assessment of educational outcomes and the improvement of undergraduate education gains momentum, and as many efforts to establish better bridges between high schools and colleges take hold, the predictive validity of the SAT will become less of a "policy issue" as far more serious issues in academic policies and standards are addressed.

References

Donlon, Thomas F. (Ed.). The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests. New York: College Entrance Examination Board, 1984.

Fincher, Cameron. "Is the SAT Worth Its Salt? An Evaluation of the Use of the Scholastic Aptitude Test in the University System of Georgia over a Thirteen-Year Period." Review of Educational Research, 1974, 44, 293-305.

______________. "Differential Validity and Test Bias." Personnel Psychology, 1965, 28, 481-500.

______________. Adult Learners and the SAT in the University System of Georgia. Athens: Institute of Higher Education, University of Georgia, 1984.

______________. Uses of the SAT in the University System of Georgia. College Board Report No. 86-5. New York: College Entrance Examination Board, 1986.

______________. "The Predictive Contributions of the SAT in a Statewide System of Public Higher Education." In Measures in the College Admissions Process: A College Board Colloquium. New York: College Entrance Examination Board, 1986.

On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977.

APPENDIX A

FIVE-YEAR AVERAGES IN VALIDITY COEFFICIENTS FOR HIGH SCHOOL AVERAGES AND SAT SCORES

Figure UN-1.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: University of Georgia Freshmen 1958-1987

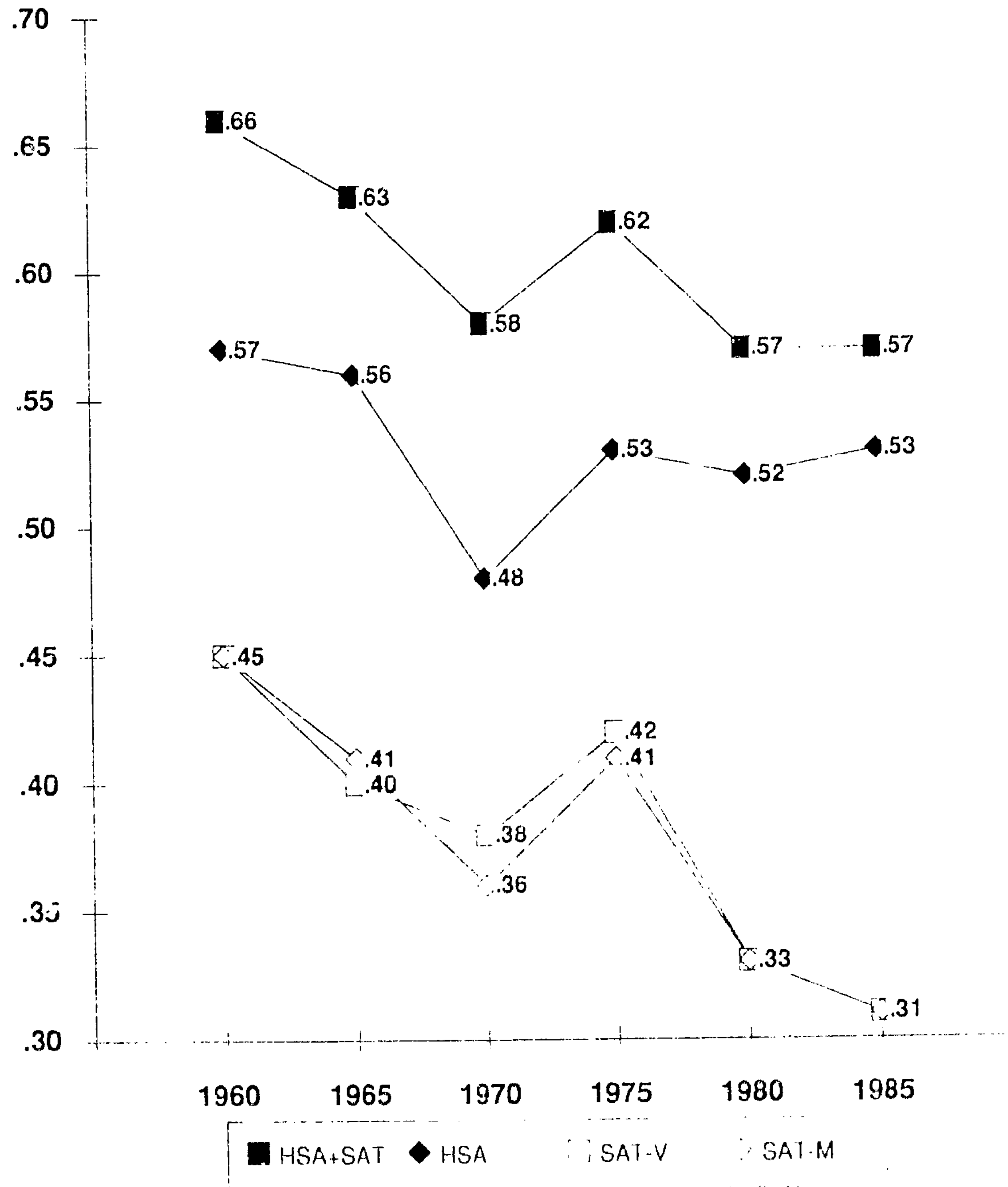


Figure UN-2.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Georgia State University Freshmen 1958-1987

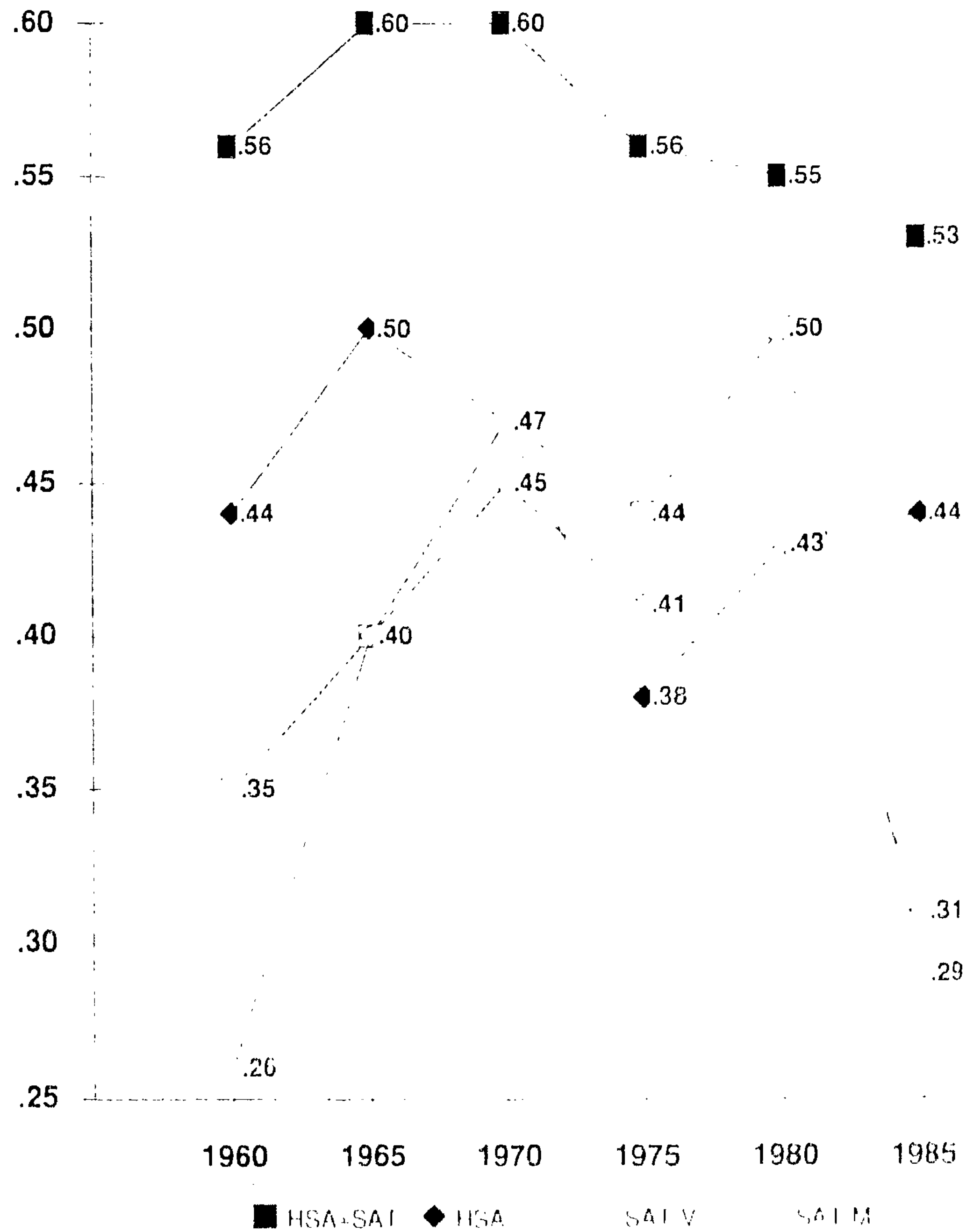


Figure UN-3.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Georgia Tech Freshmen 1958-1987

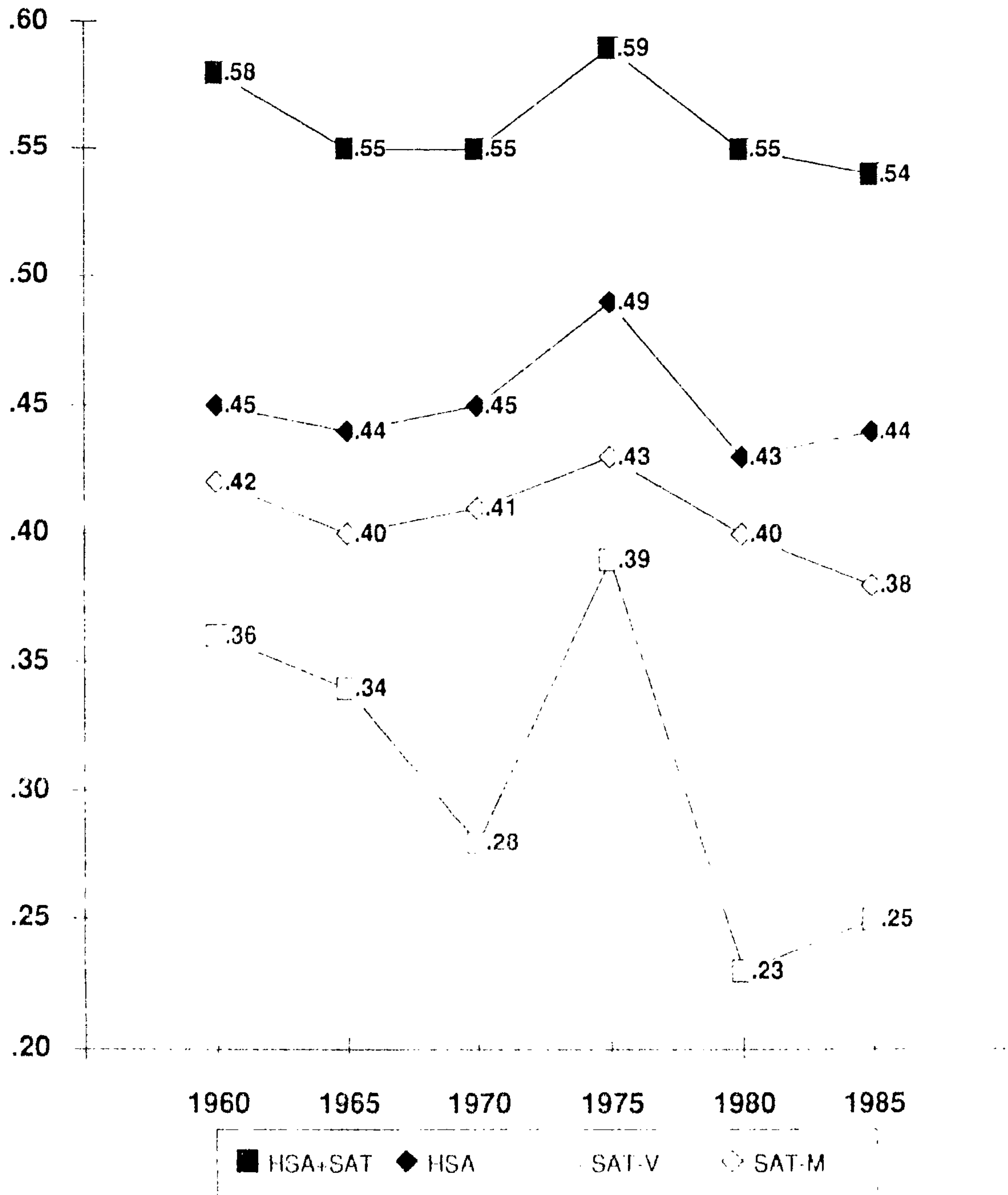


Figure Four-1.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Armstrong State College Freshmen 1958-1987

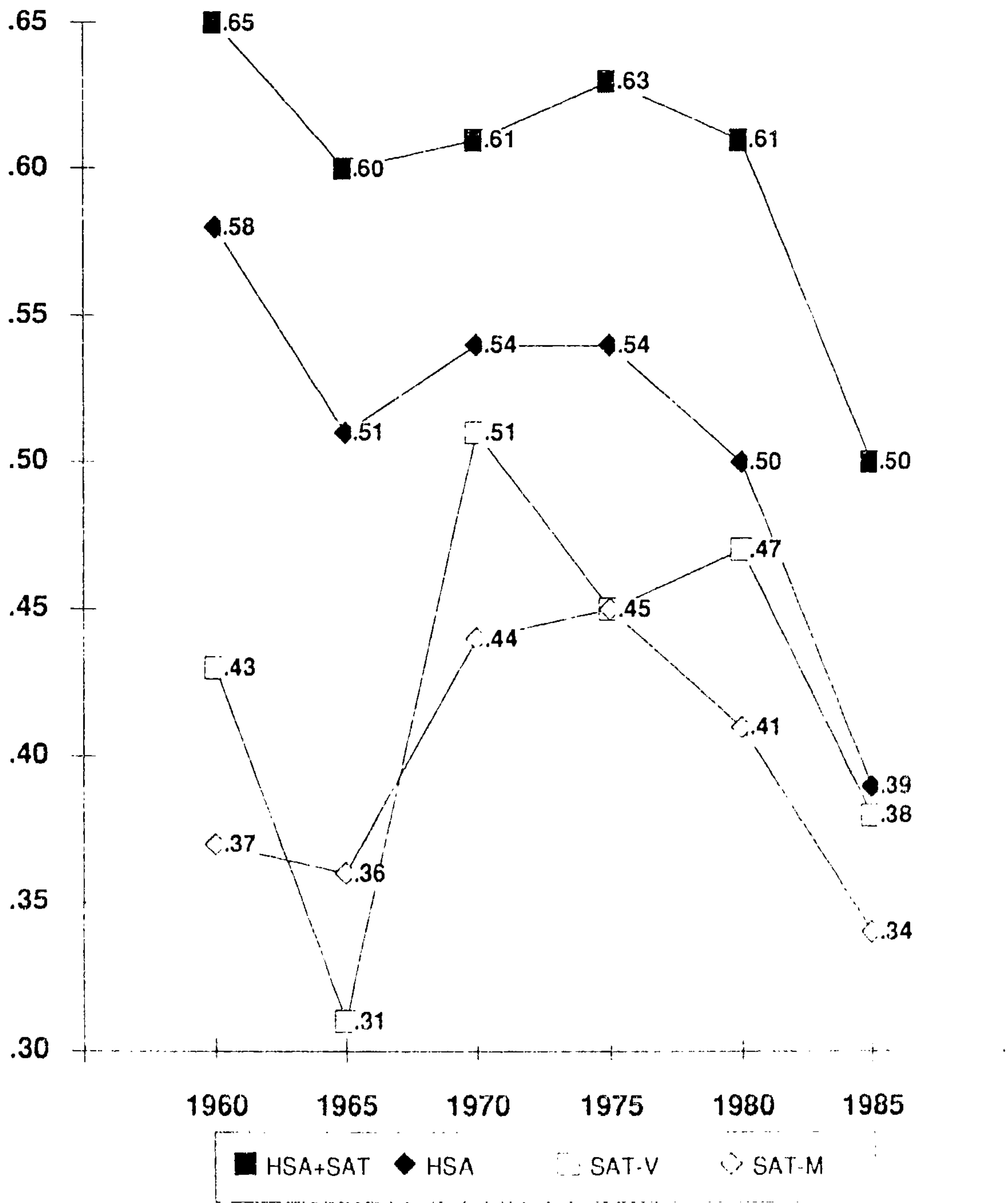


Figure Four-2.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Augusta College Freshmen 1958-1987

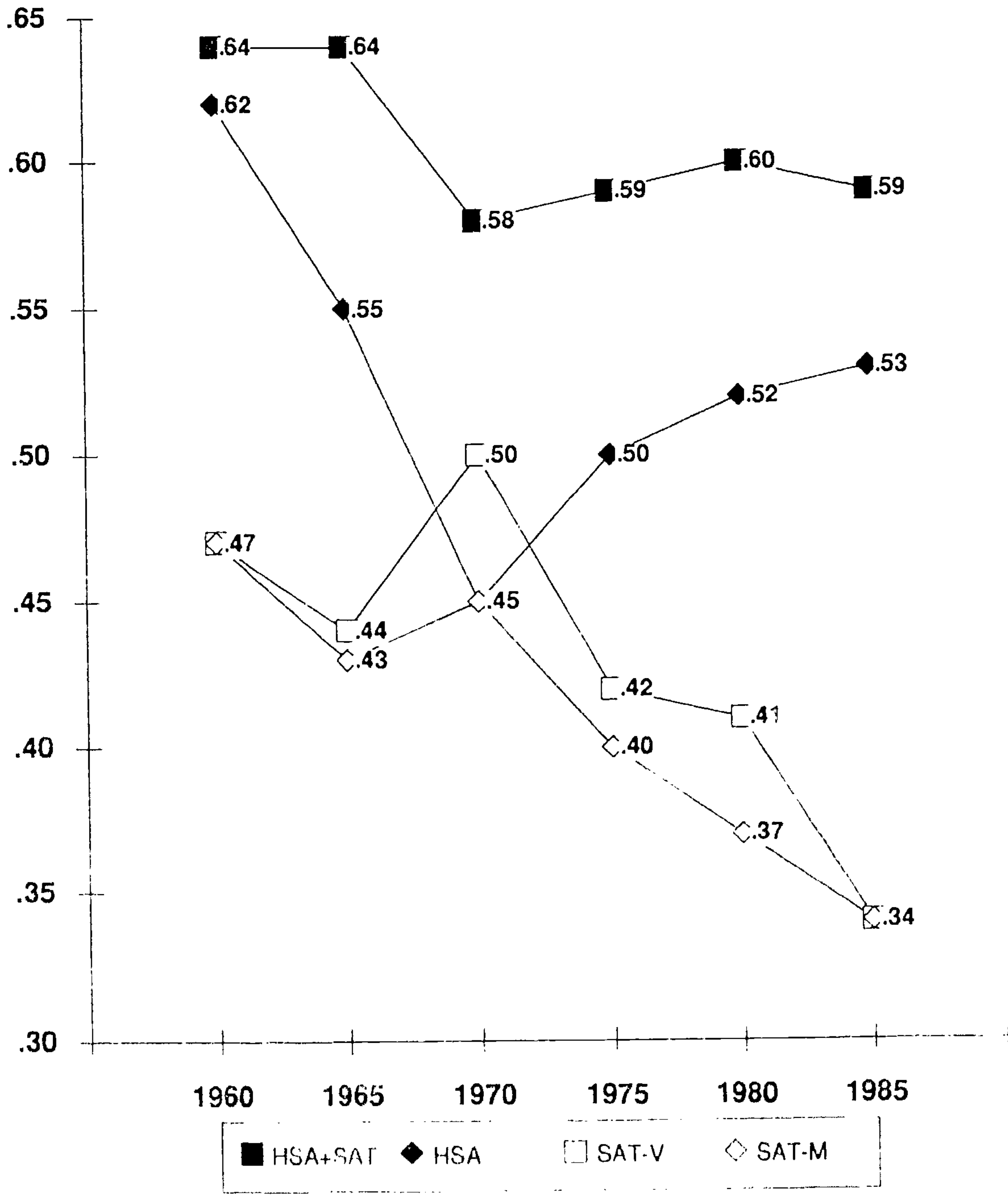


Figure Four-3.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
Columbus College Freshmen 1958-1987

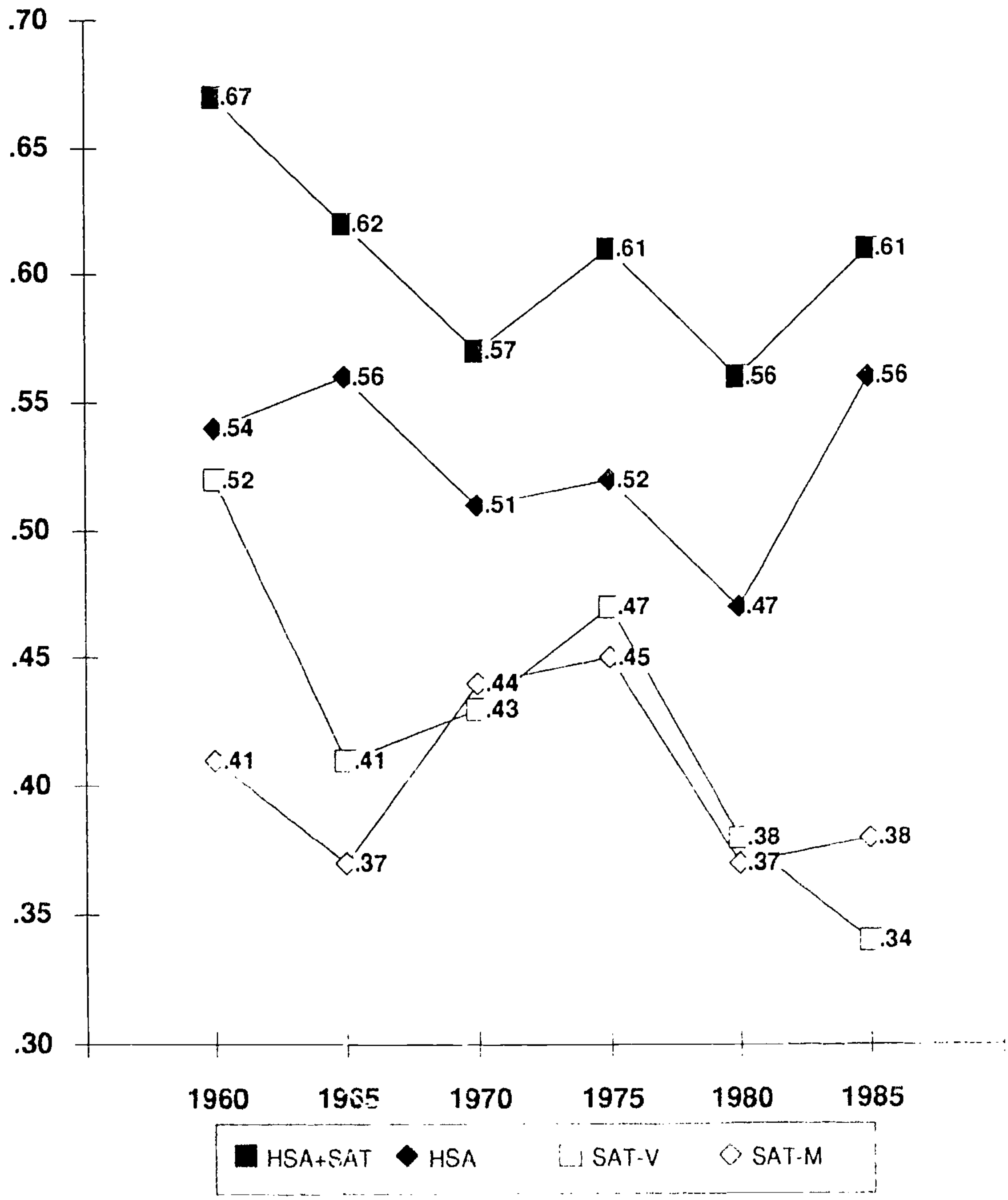


Figure Four-4.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Georgia College Freshmen 1958-1987

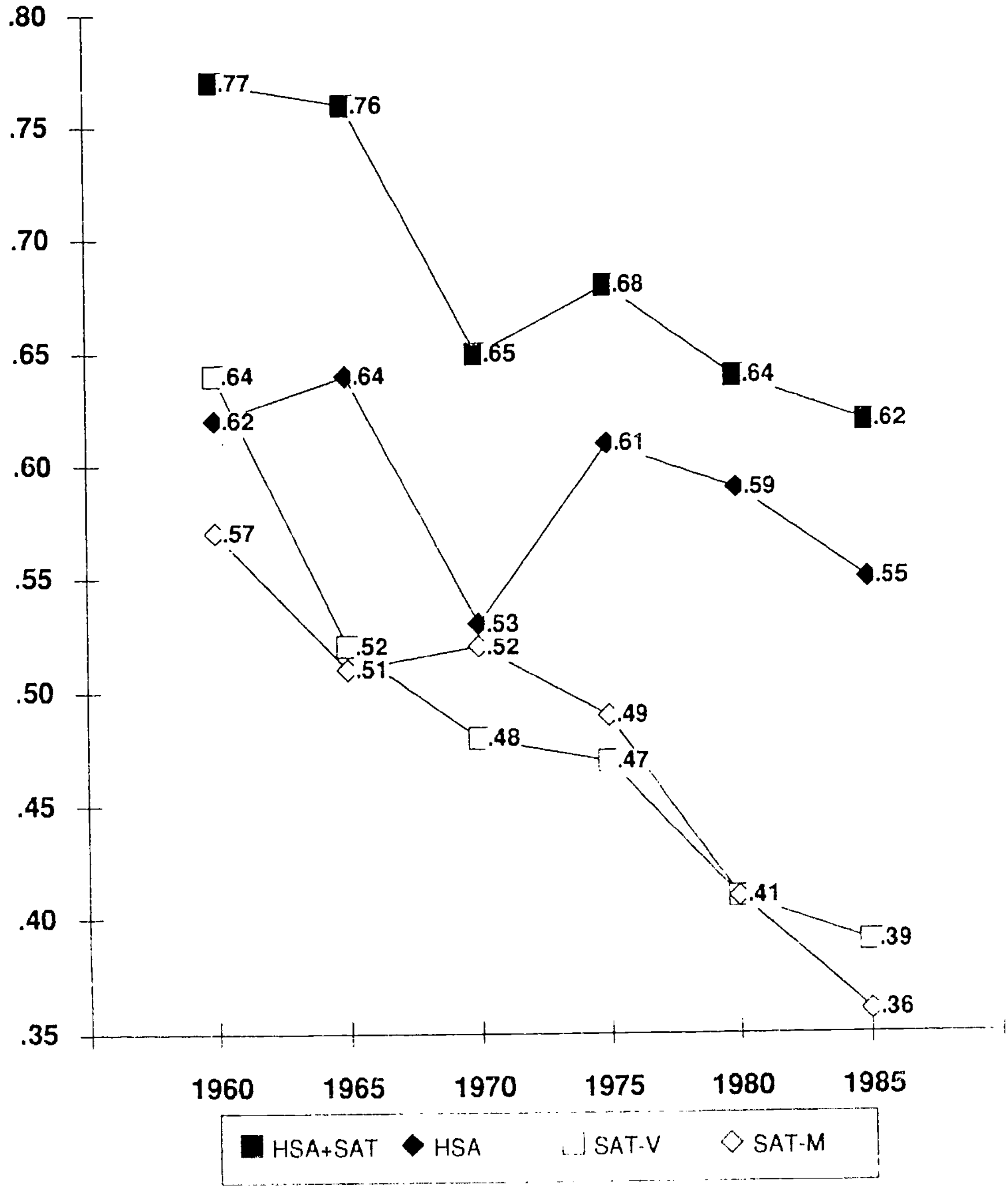


Figure Four-5.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Georgia Southern College Freshmen 1958-1987

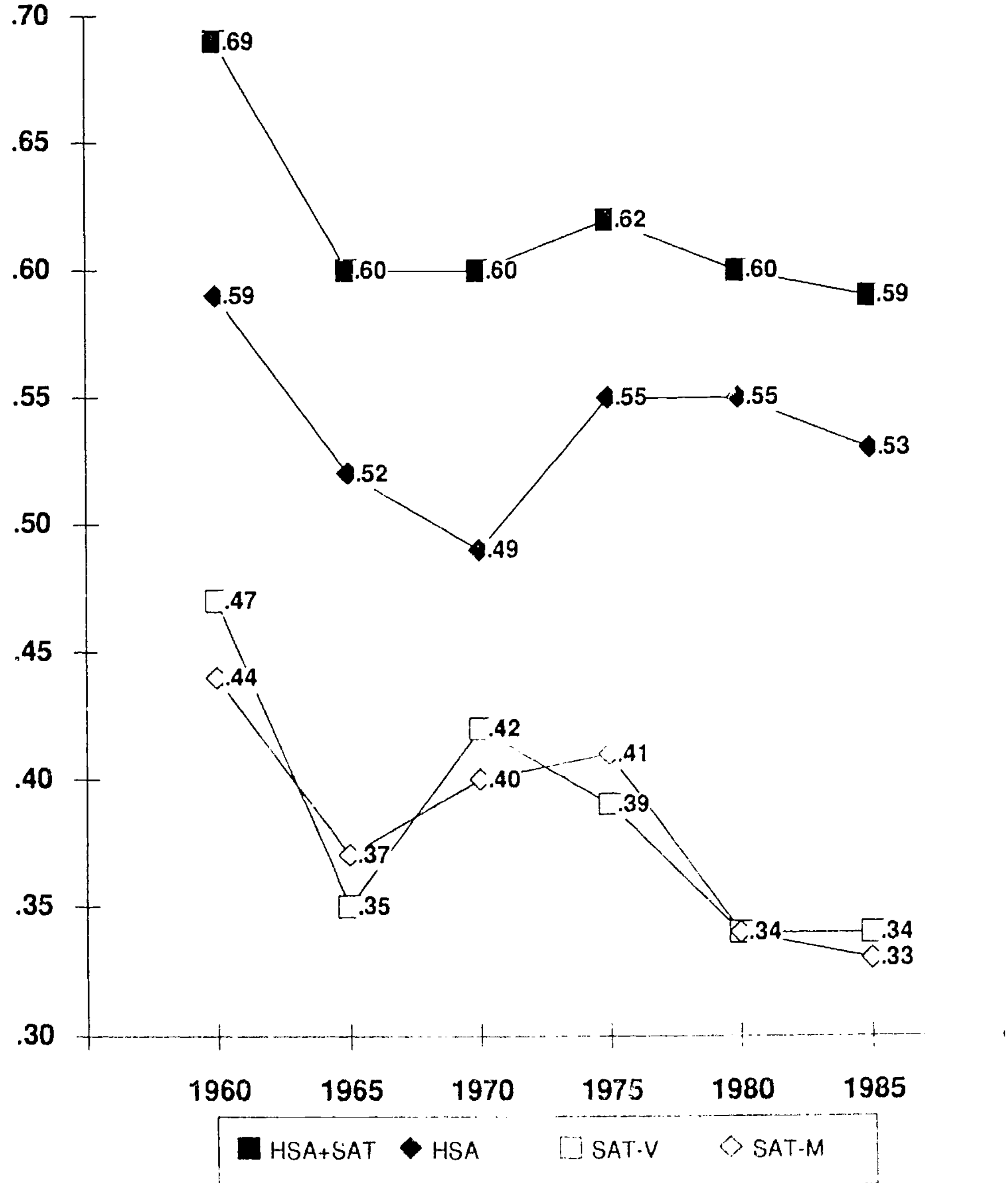


Figure Four-6.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
Georgia Southwestern College Freshmen 1958-1987

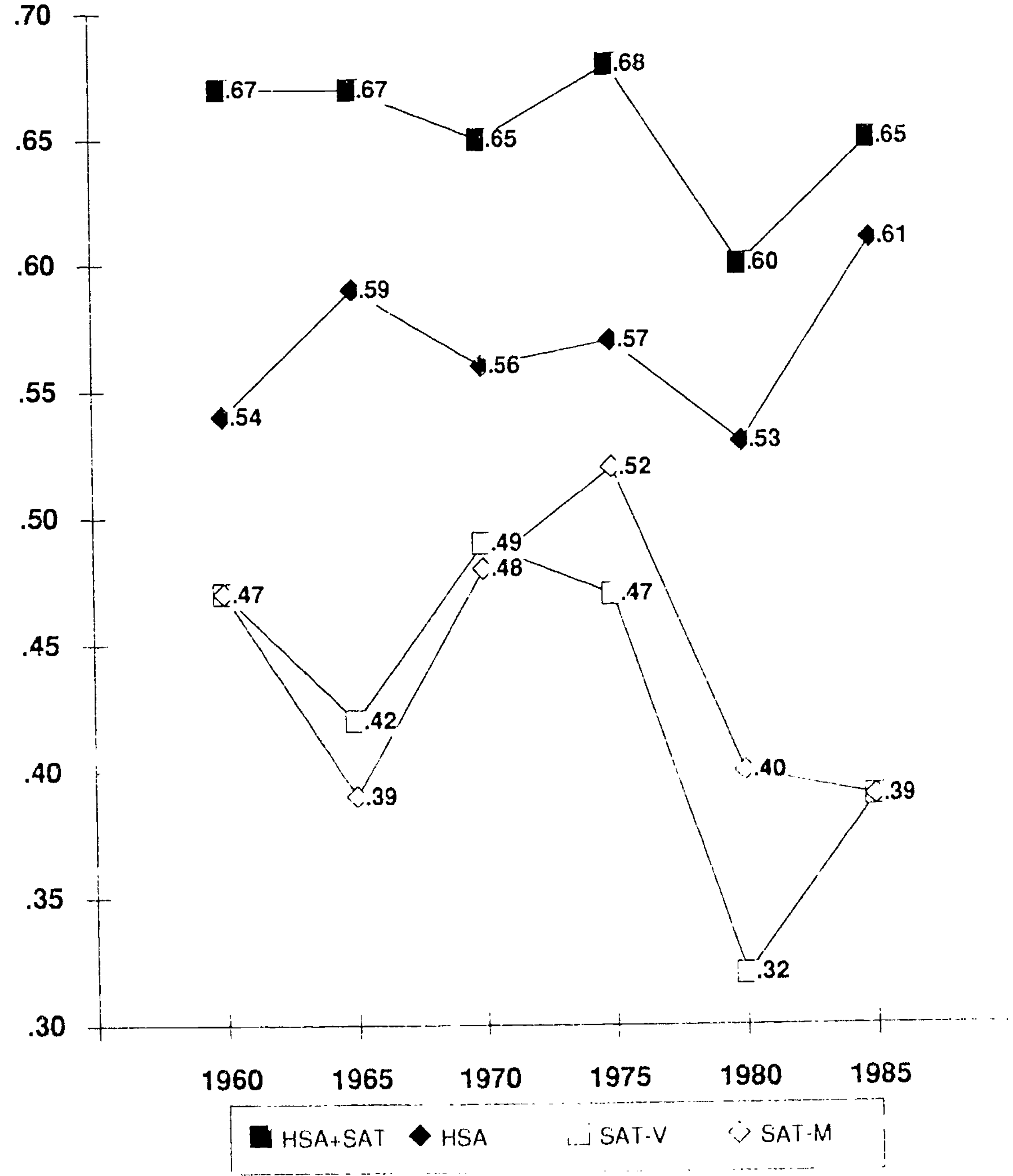


Figure Four-7.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Kennesaw College Freshmen 1963-1987

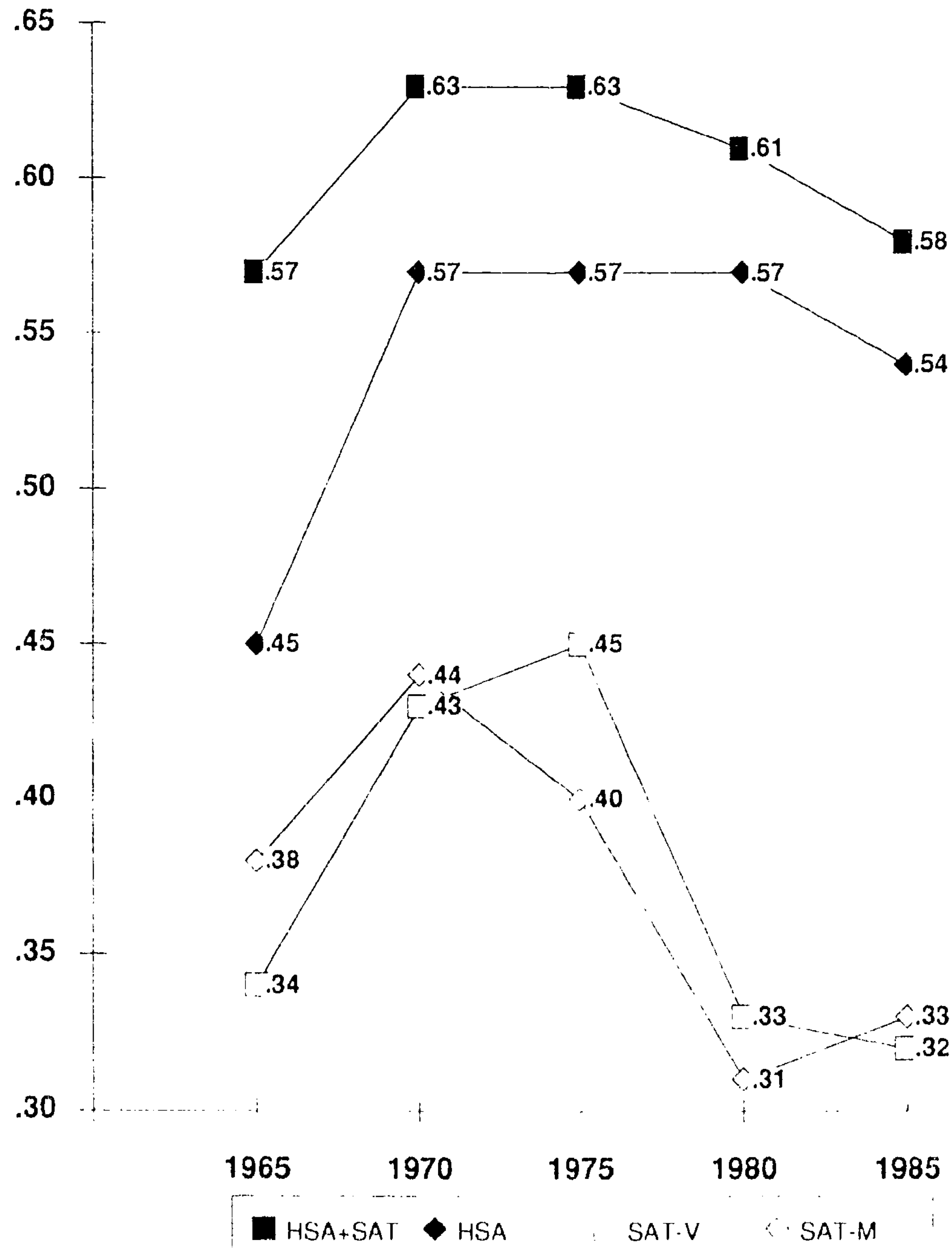


Figure Four-8.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
North Georgia College Freshmen 1958-1987

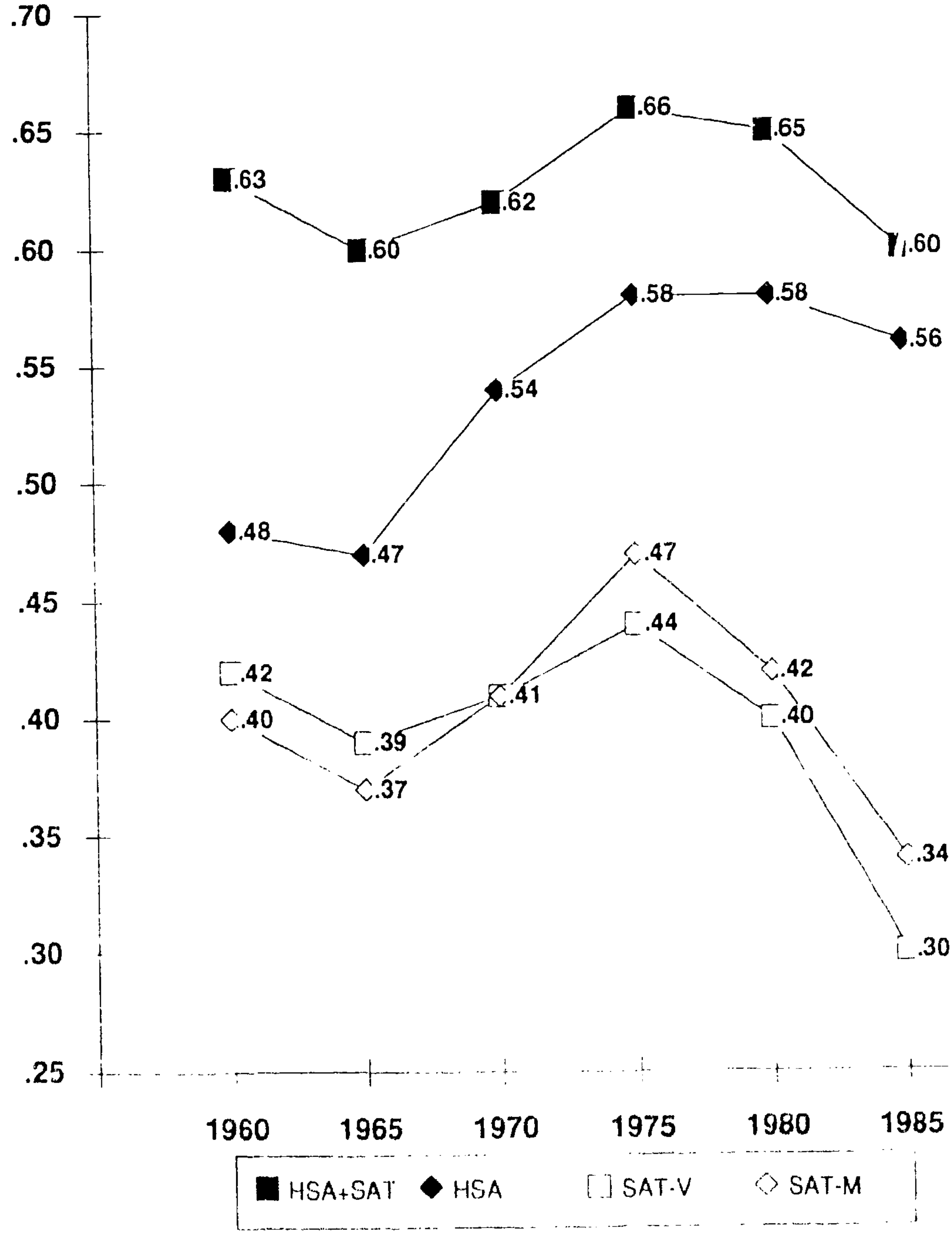


Figure Four-9.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Valdosta State College Freshmen 1958-1987

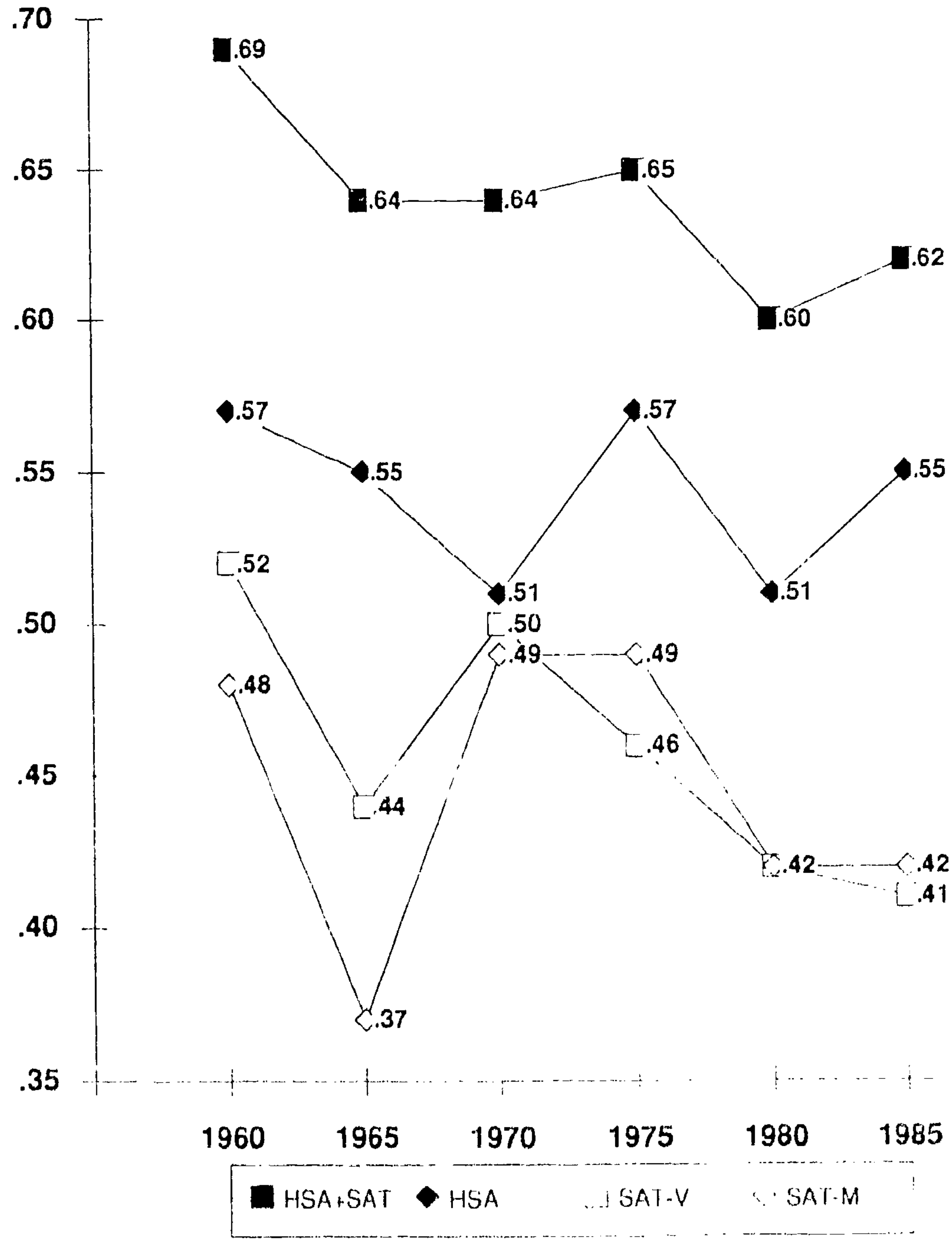


Figure Four-10.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: West Georgia College Freshmen 1958-1987

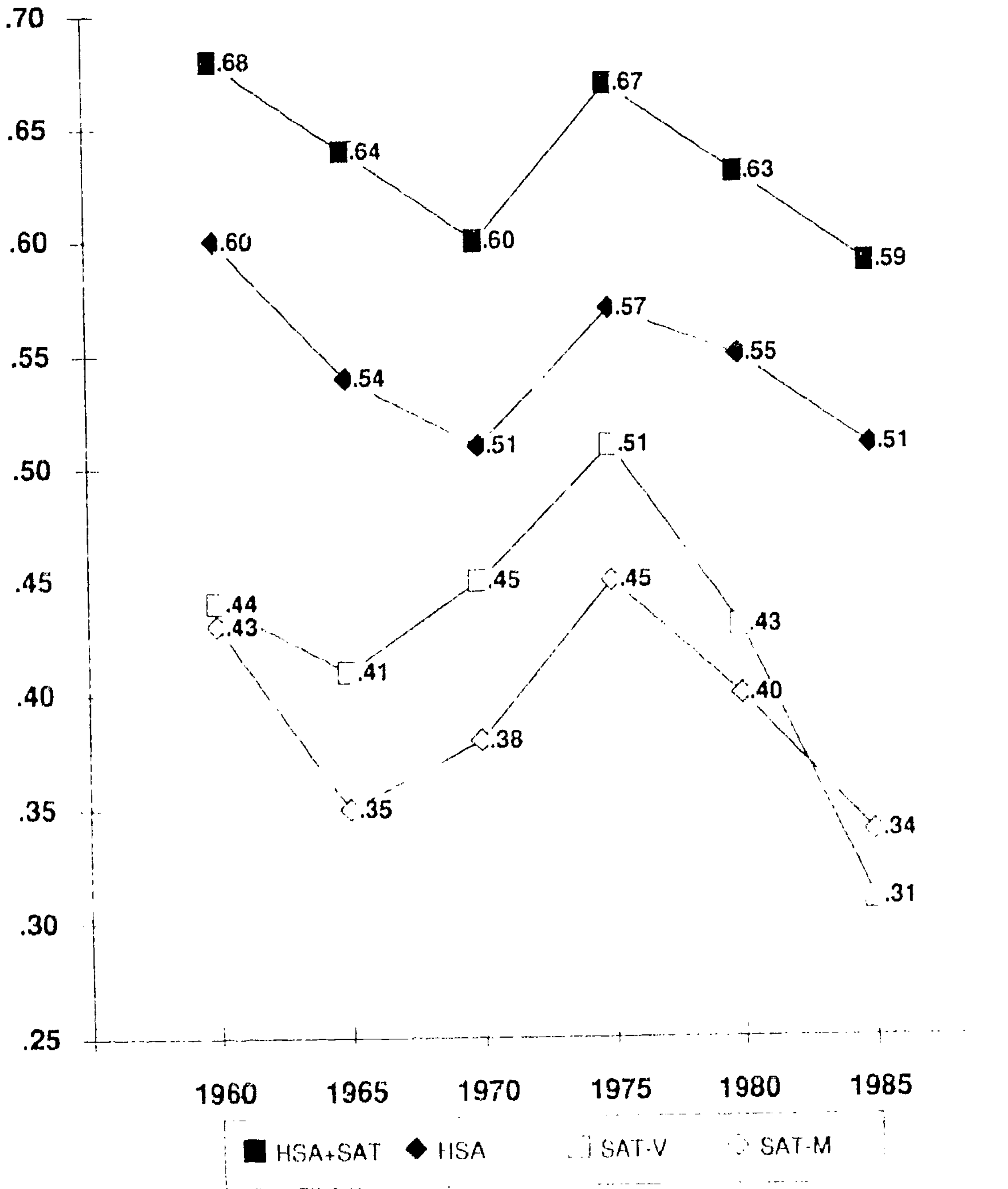


Figure Four-11.

Five-Year Averages in Validity Coefficients for High School Av:rages and SAT Scores: Southern Technical Institute Freshmen 1958-1987

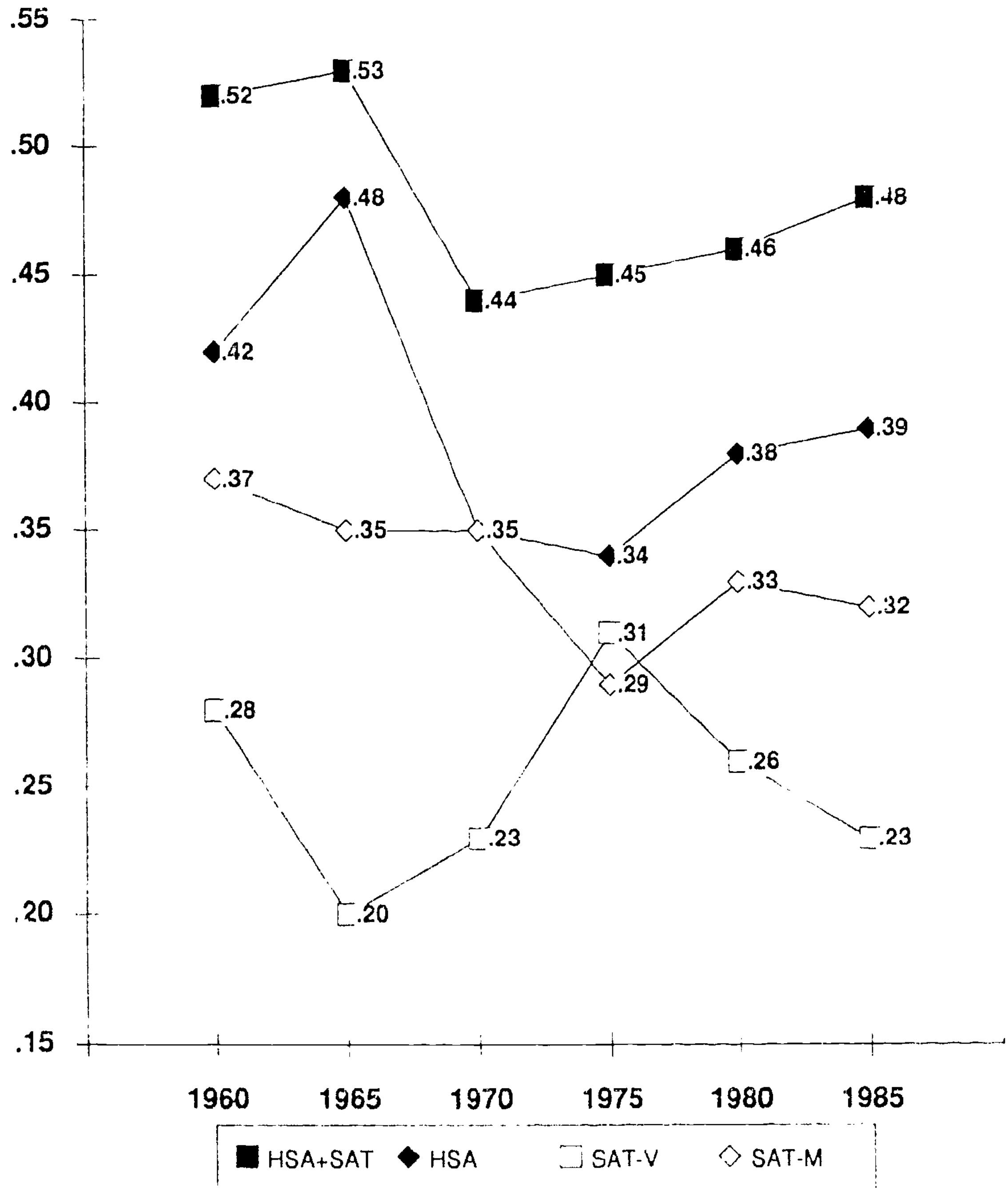


Figure Four-12.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Clayton State College Freshmen 1968-1987

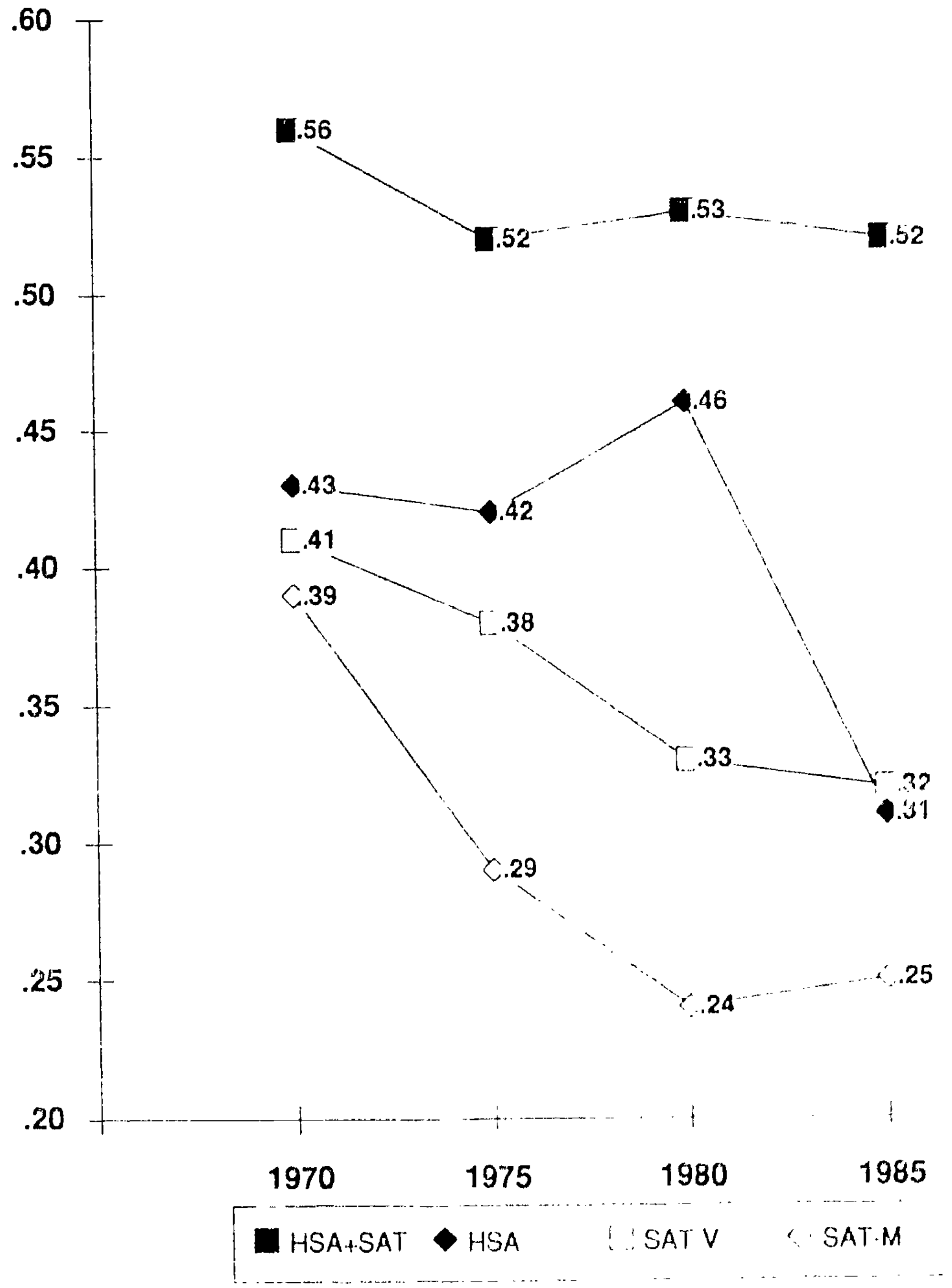


Figure Two-1.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
Abraham Baldwin Agricultural College Freshmen 1958-1987

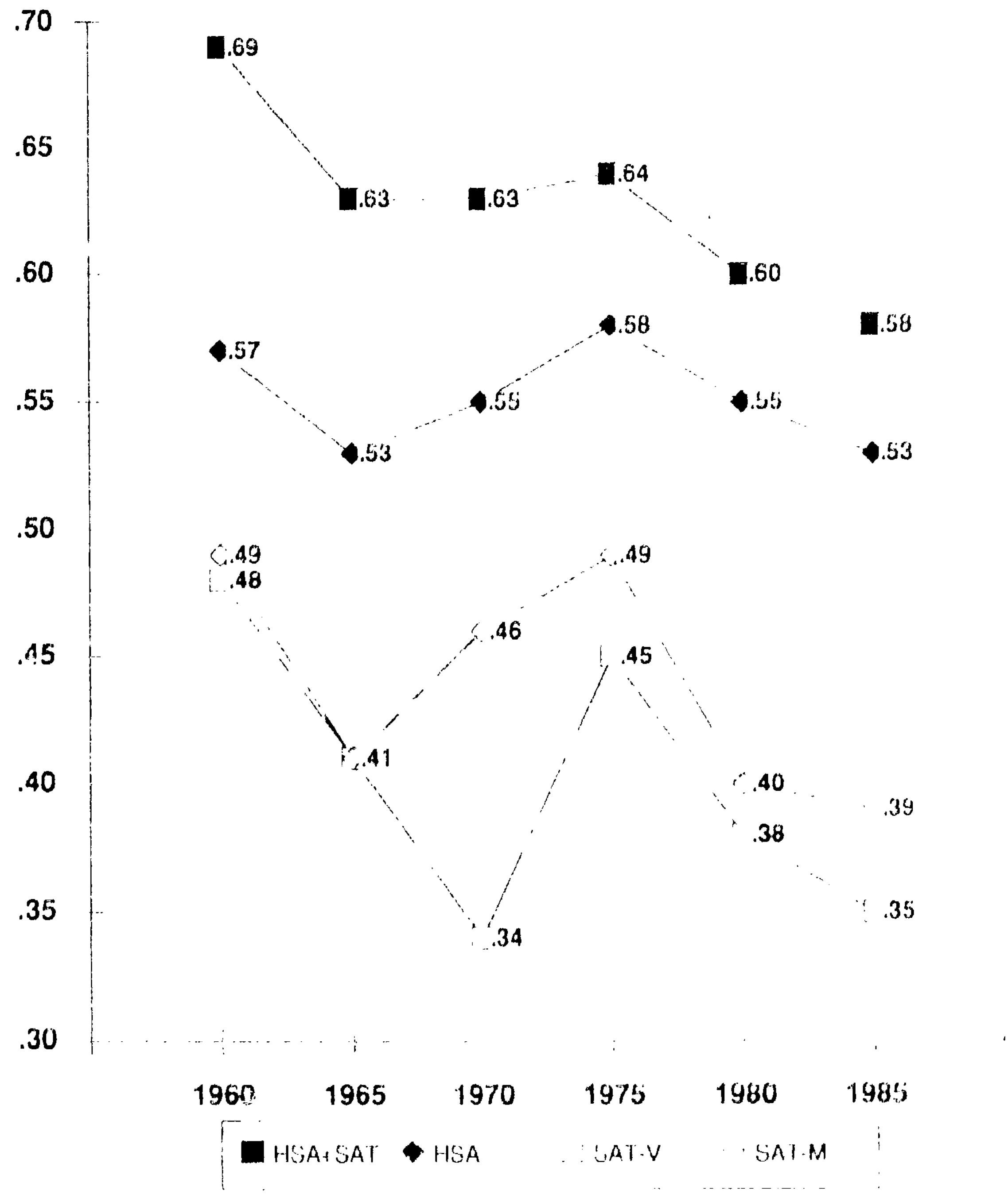


Figure Two-2.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Middle Georgia College Freshmen 1958-1987

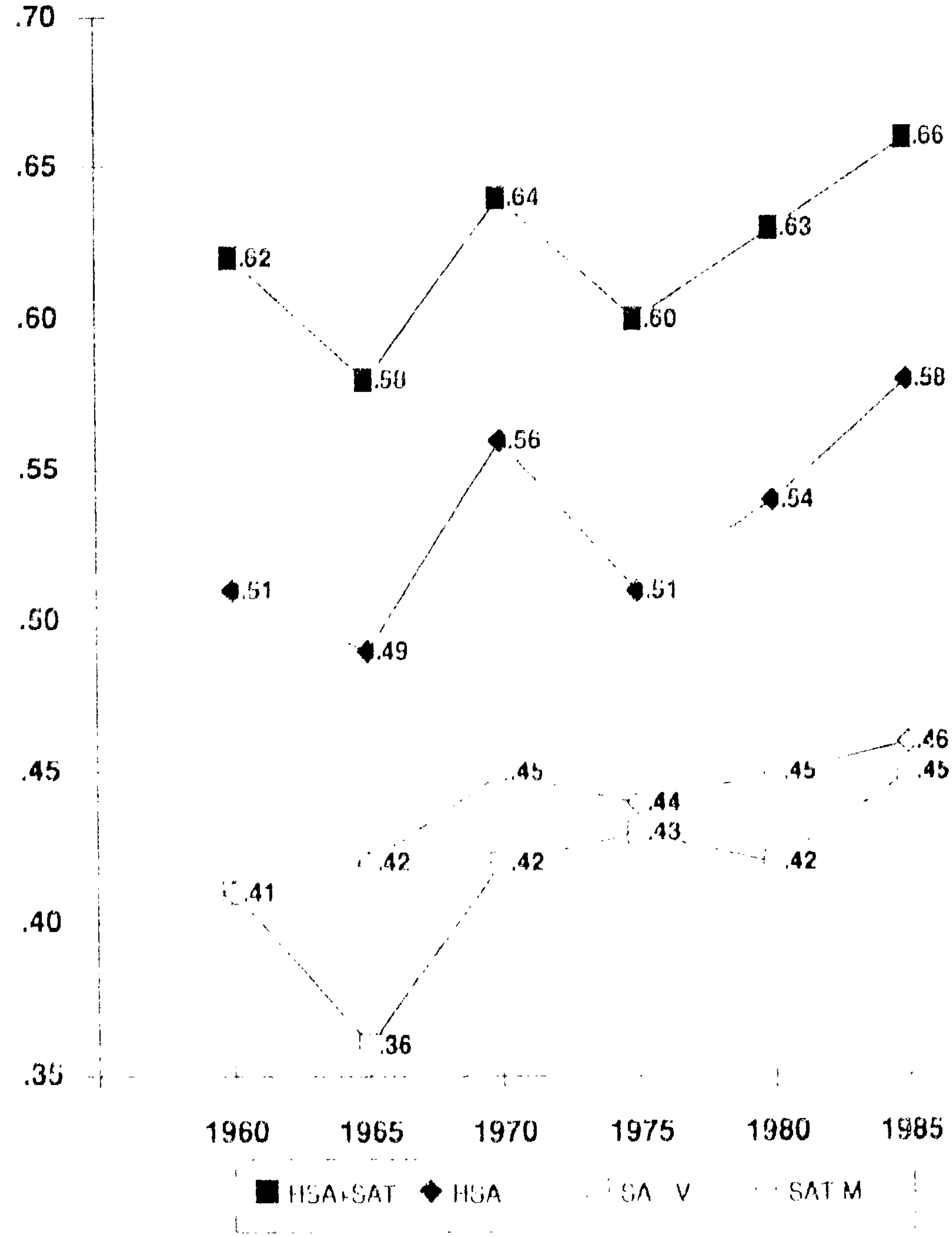


Figure Two-3.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: South Georgia College Freshmen 1958-1987

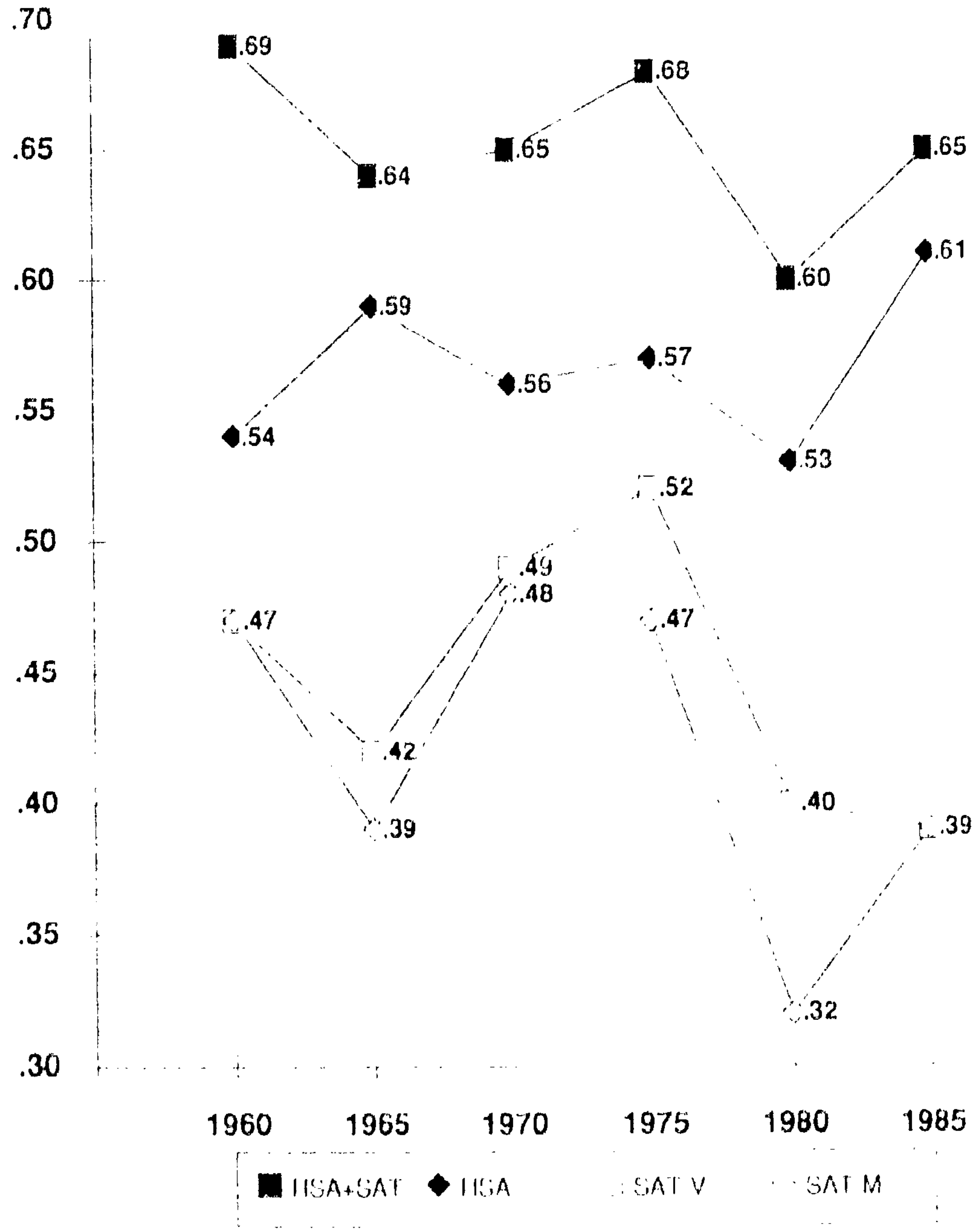


Figure Two-4.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Albany Junior College Freshmen 1963-1987

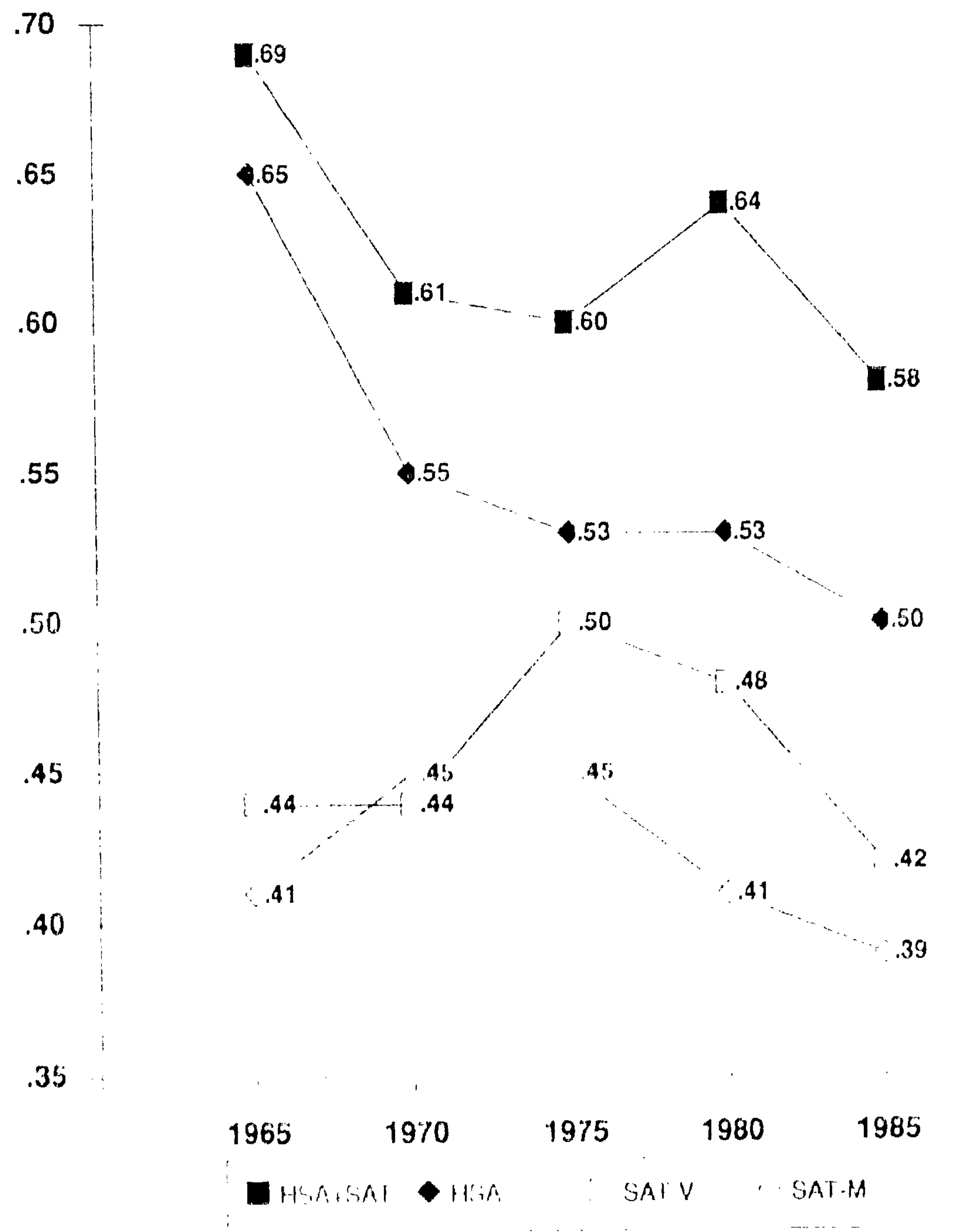


Figure Two-5.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Dalton Junior College Freshmen 1963-1987

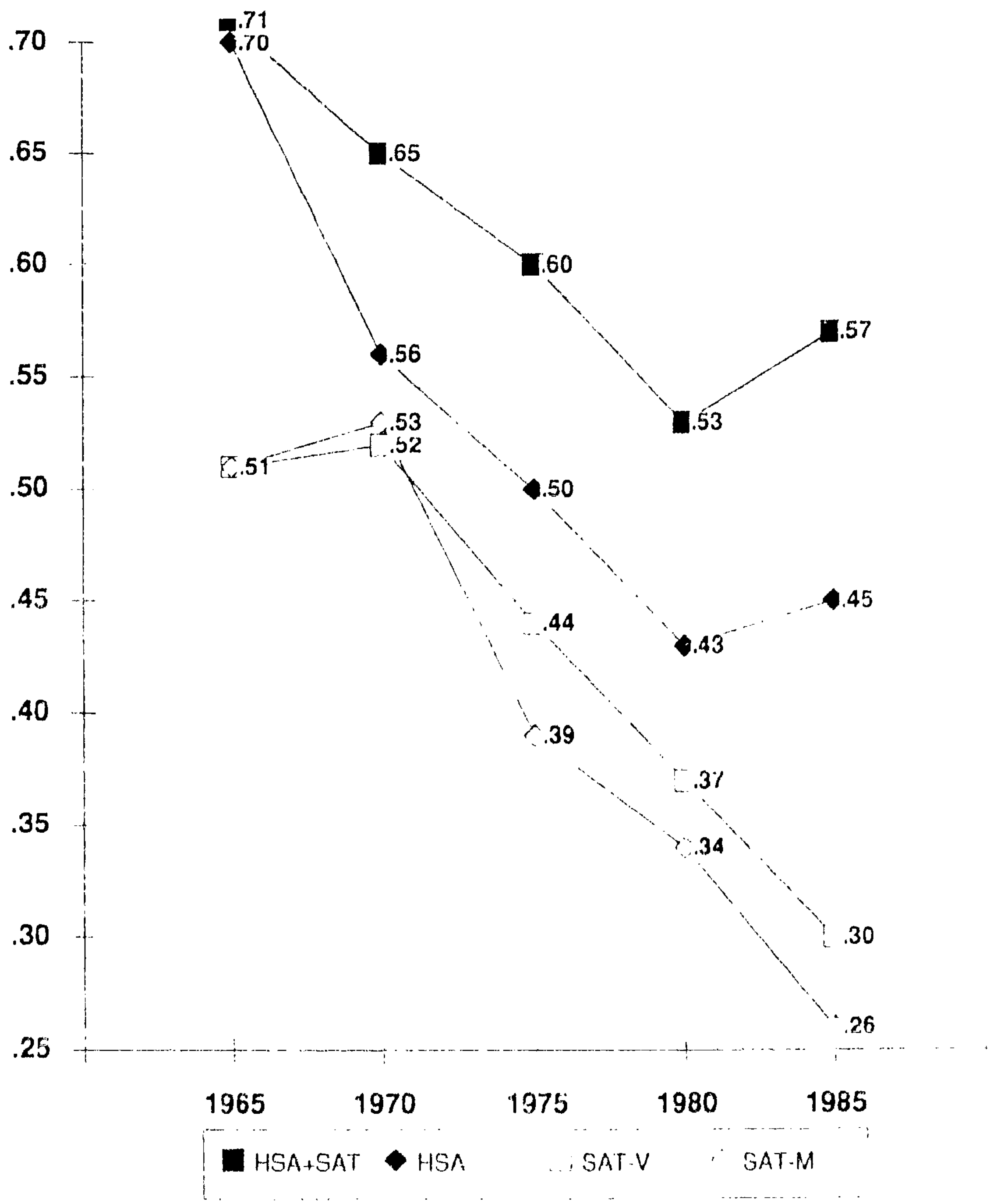


Figure Two-6.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Floyd Junior College Freshmen 1968-1987

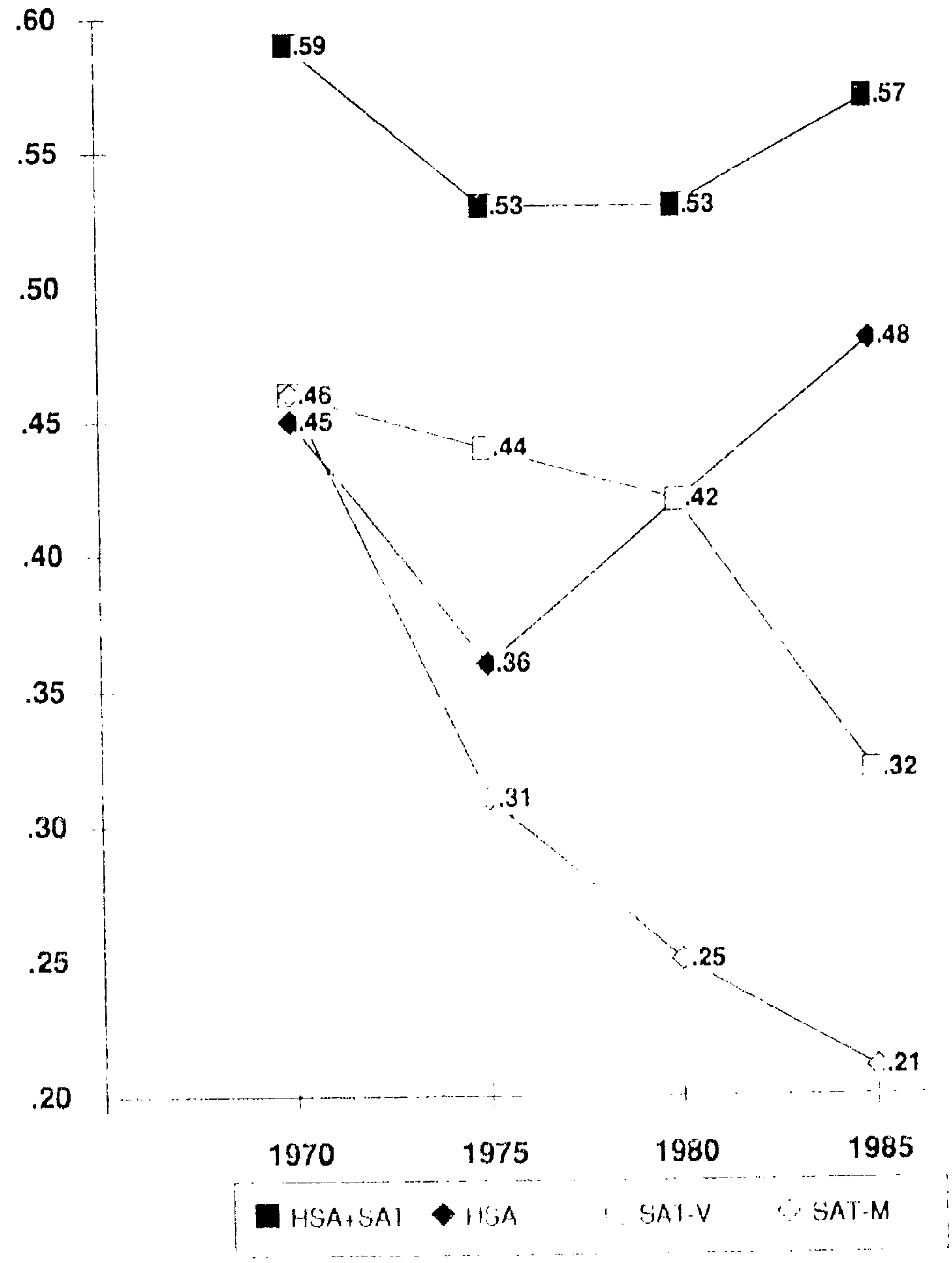


Figure Two-7.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Gainesville Junior College Freshmen 1968-1987

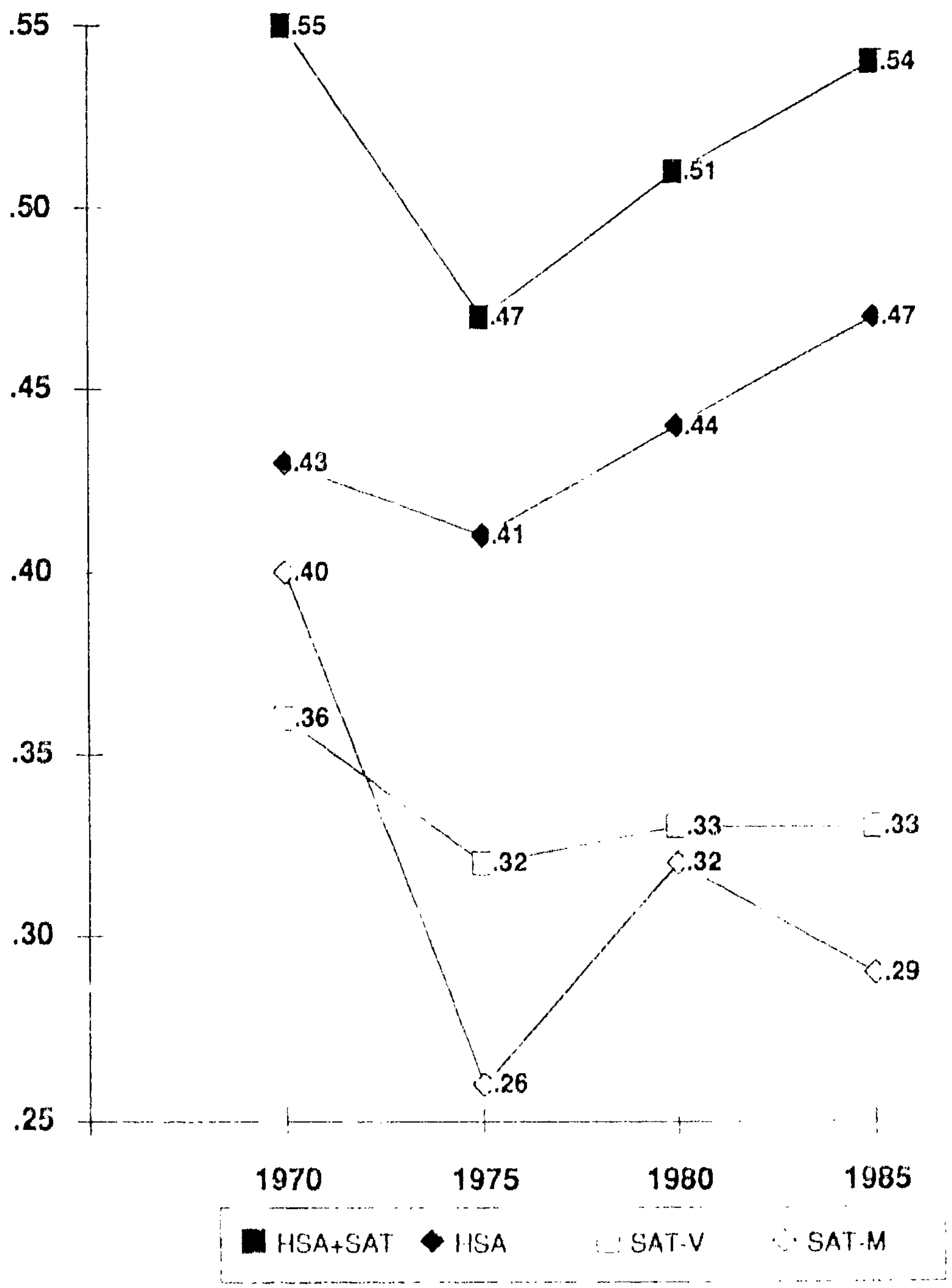


Figure Two-8.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Macon Junior College Freshmen 1968-1987

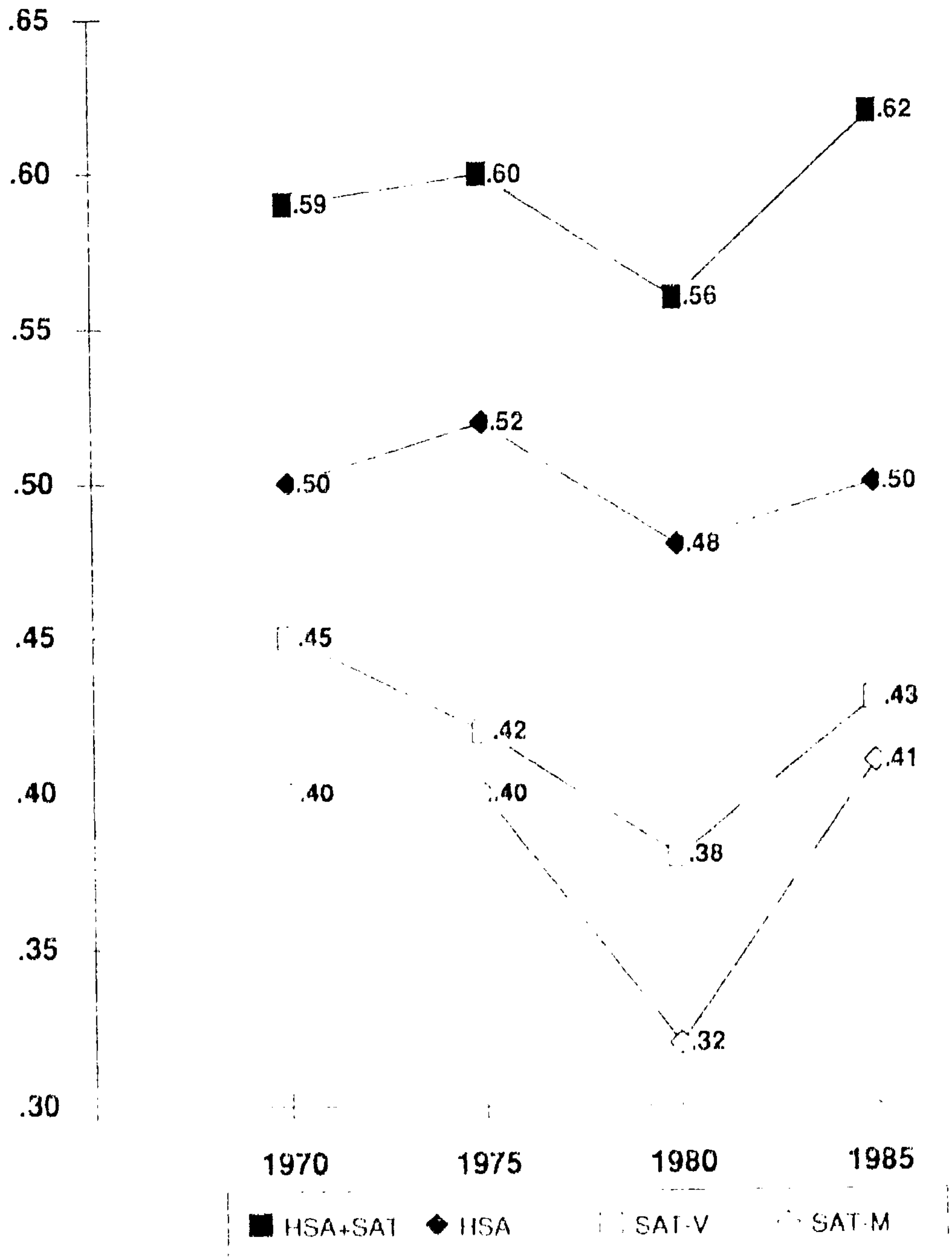


Figure Two-9.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Brunswick Junior College Freshmen 1963-1987

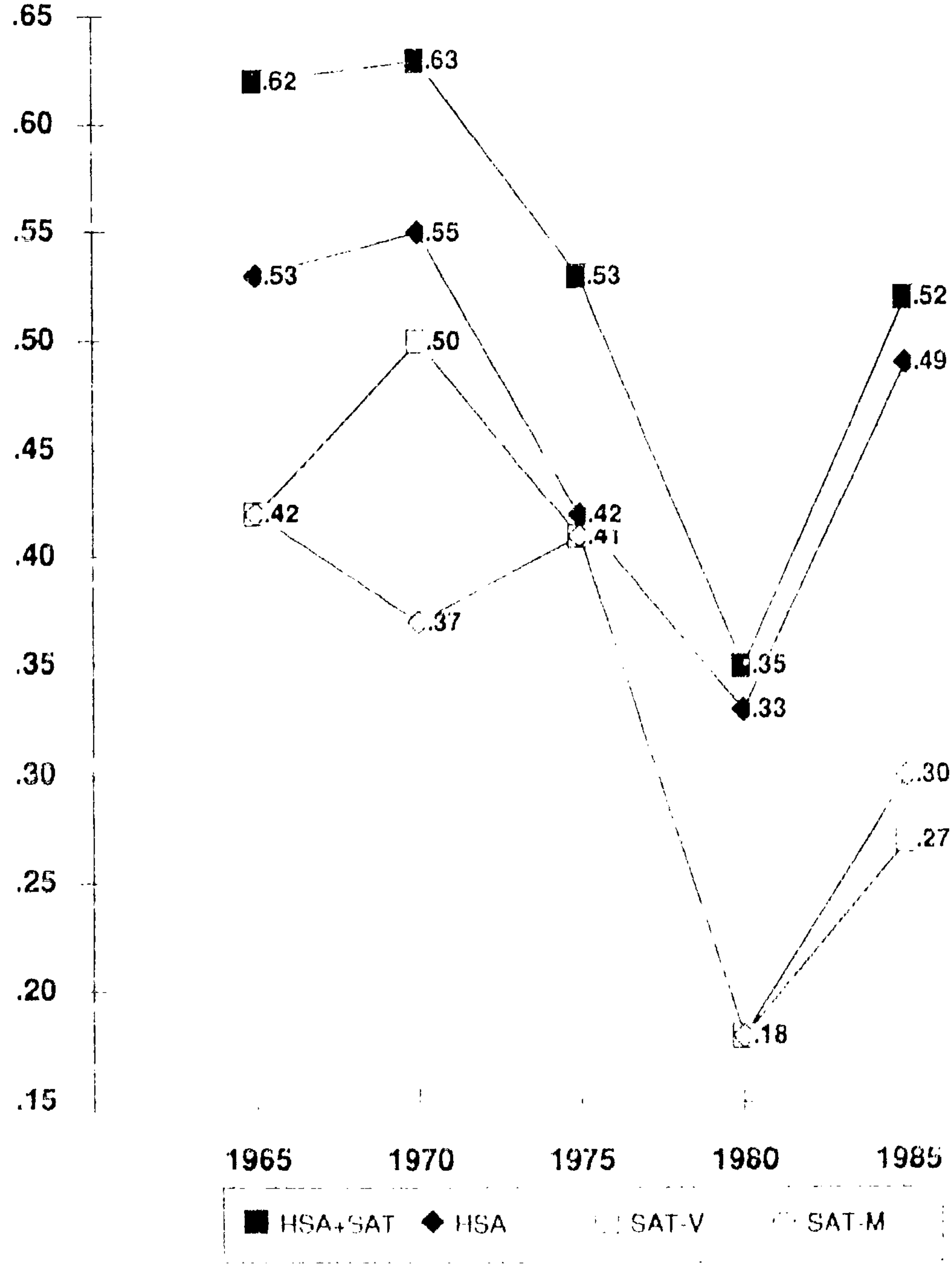


Figure HBI-1.

Five-Year Averages in Validity Coefficients for High School Averages and SAT Scores: Albany State College Freshmen 1958-1987

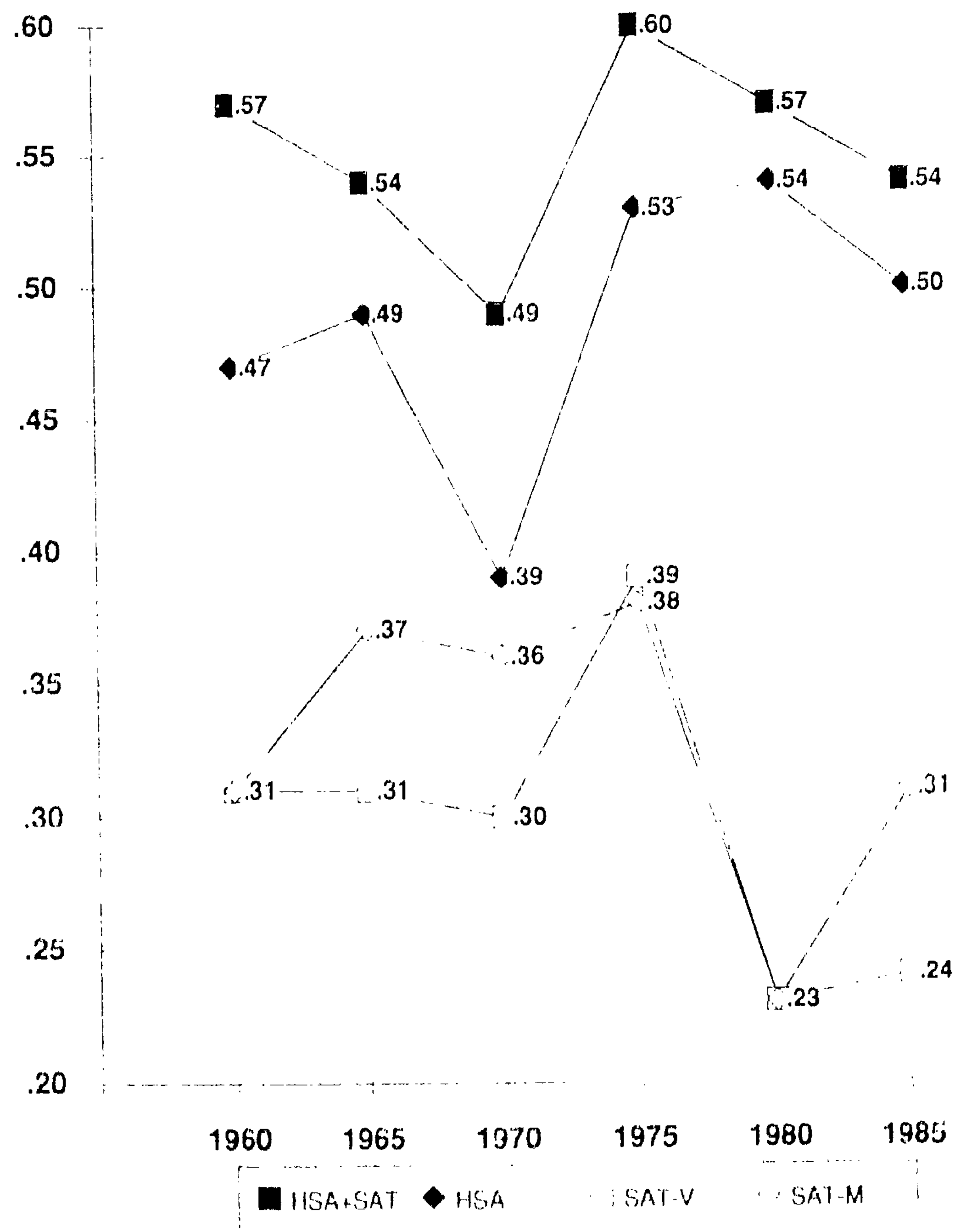


Figure HBI-2.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
Fort Valley State College Freshmen 1958-1987

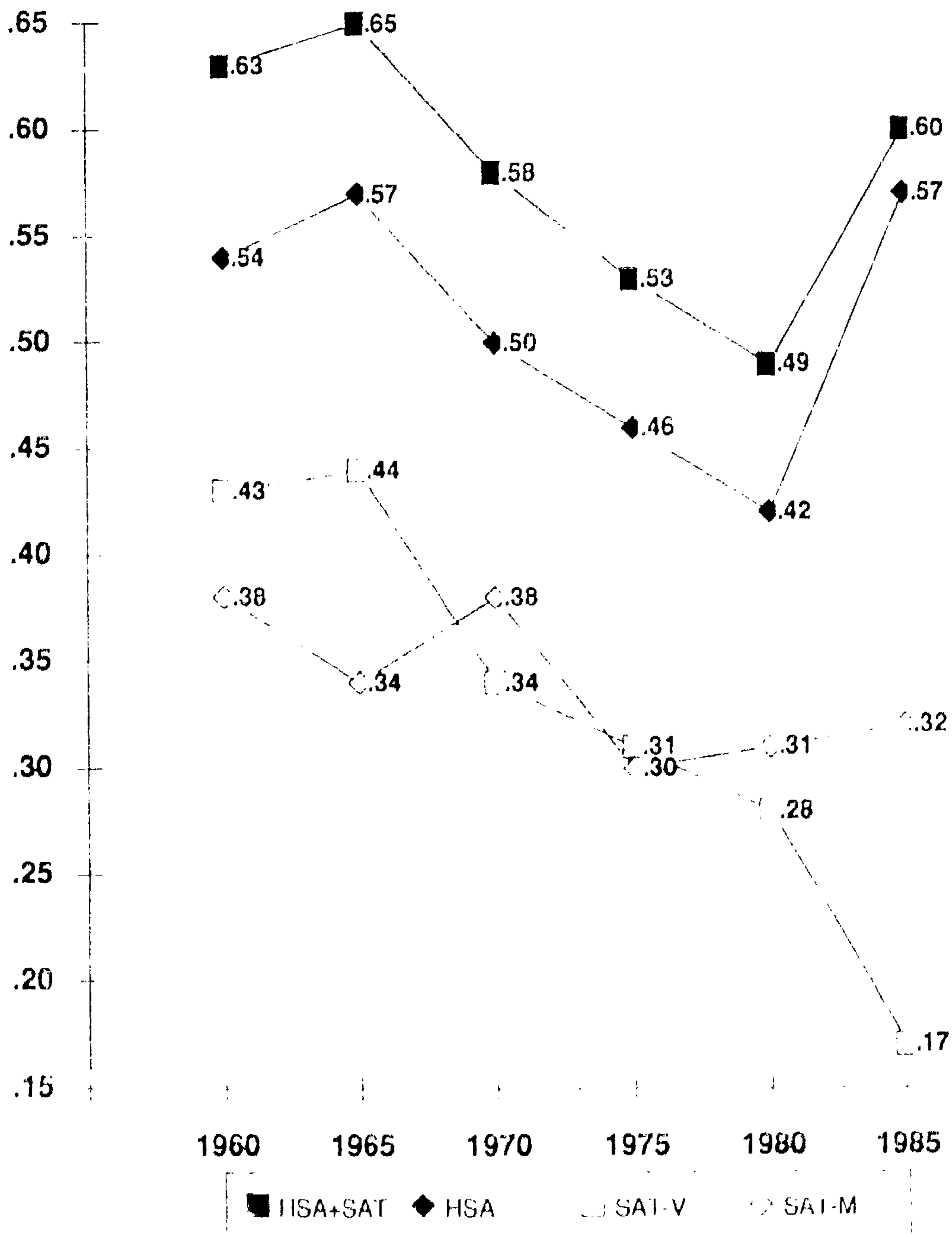
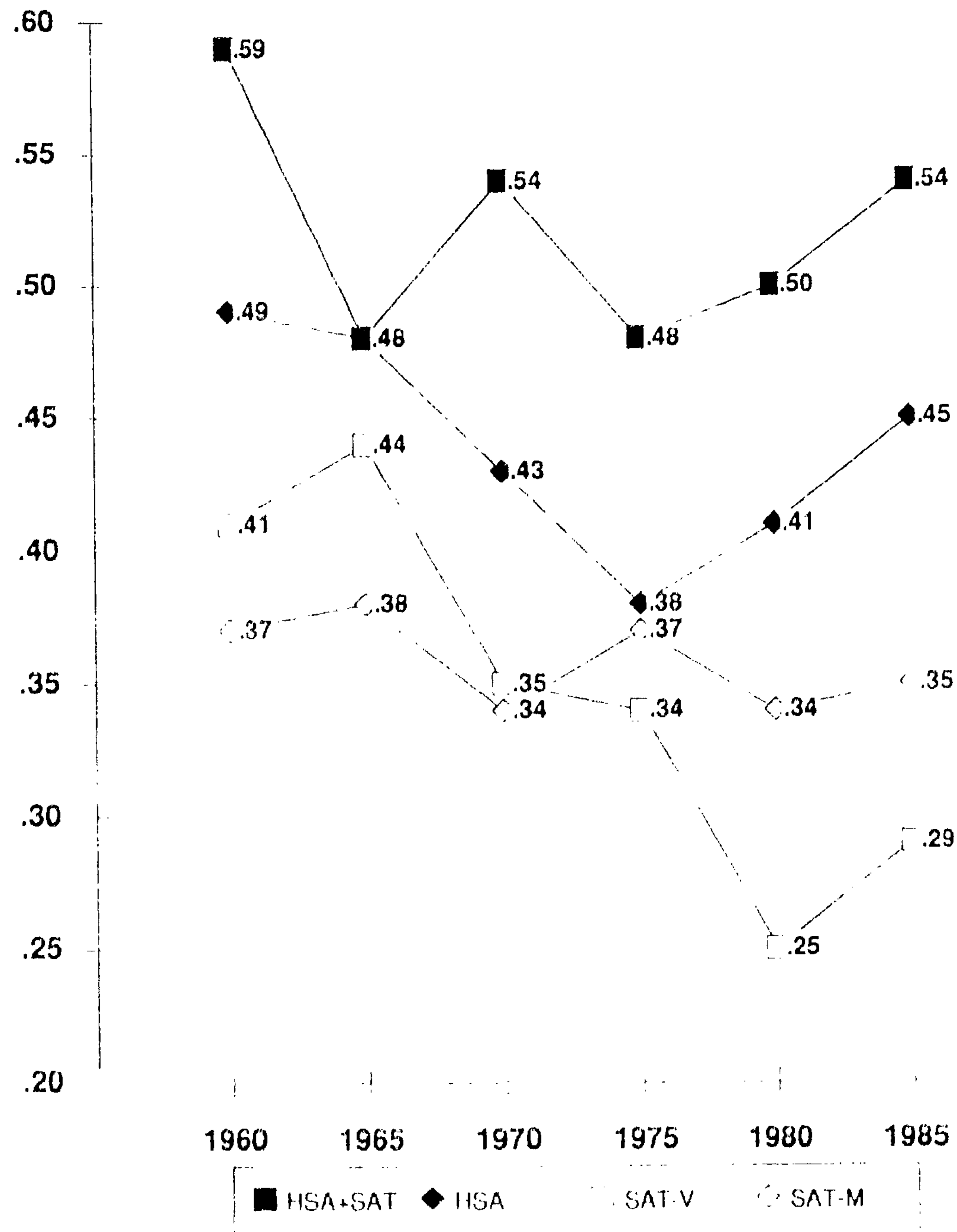


Figure HBI-3.

Five-Year Averages in Validity Coefficients
for High School Averages and SAT Scores:
Savannah State College Freshmen 1958-1987



APPENDIX B

INSTITUTIONAL PROFILES OF COOPERATING INSTITUTIONS

Figure CI-1.

Correlations of HSA and SAT with FGPA
in an Eastern Private Research University

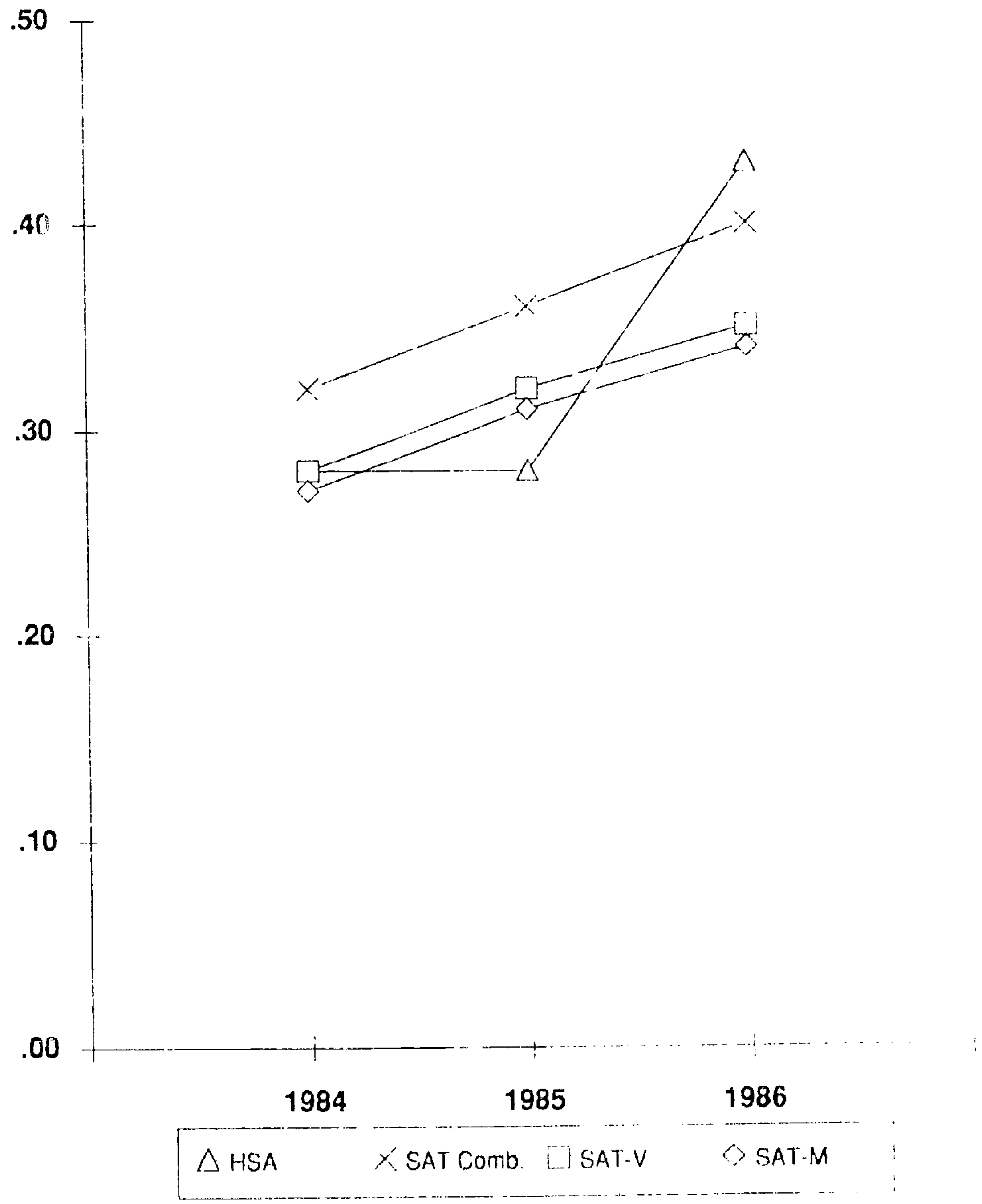


Figure CI-2A.

Correlations of HSR, SAT and HSR+SAT with FGPA for Men in an Eastern Private Research University

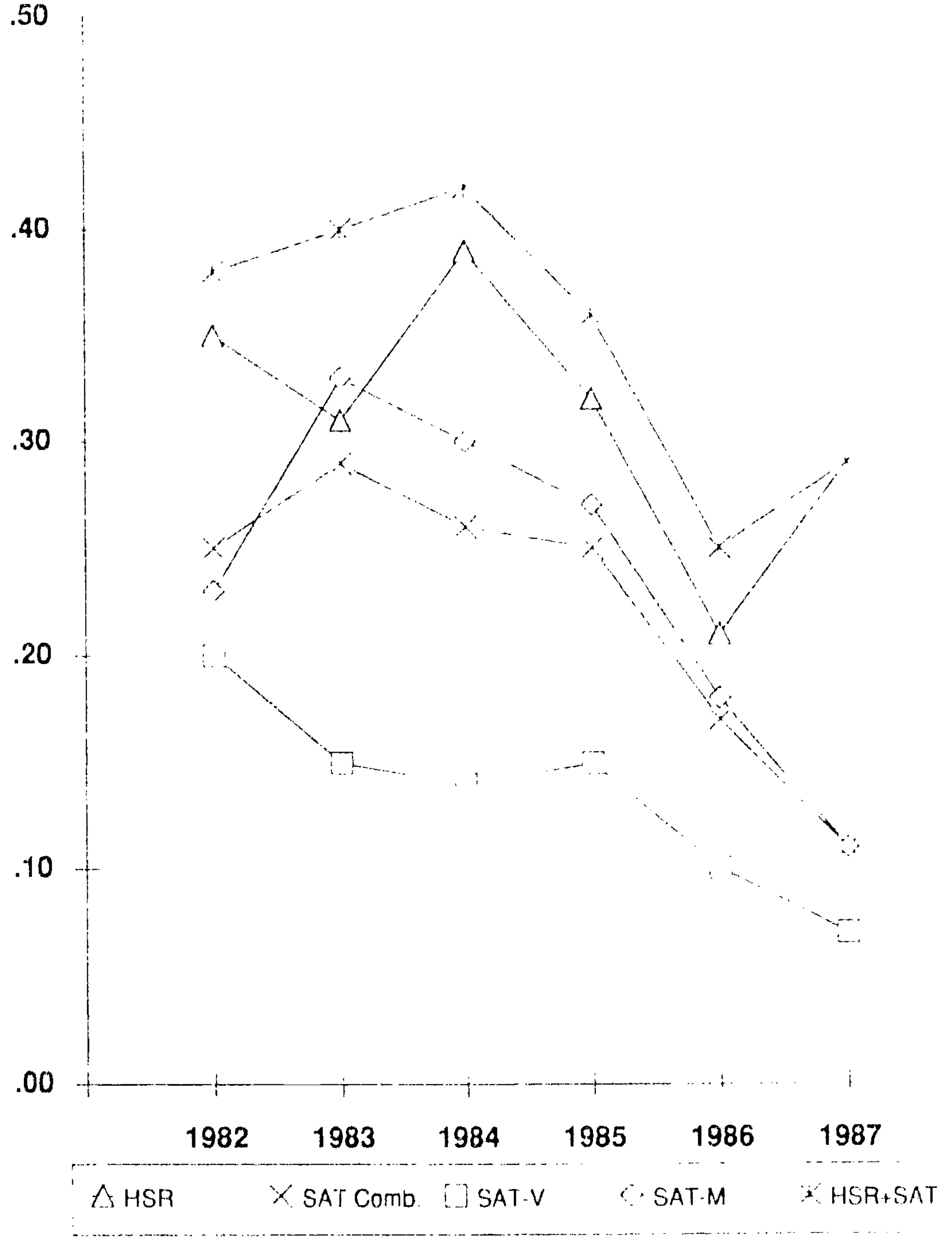


Figure CI-2B.

Correlations of HSR, SAT and HSR+SAT with FGPA
for Women in an Eastern Private Research University

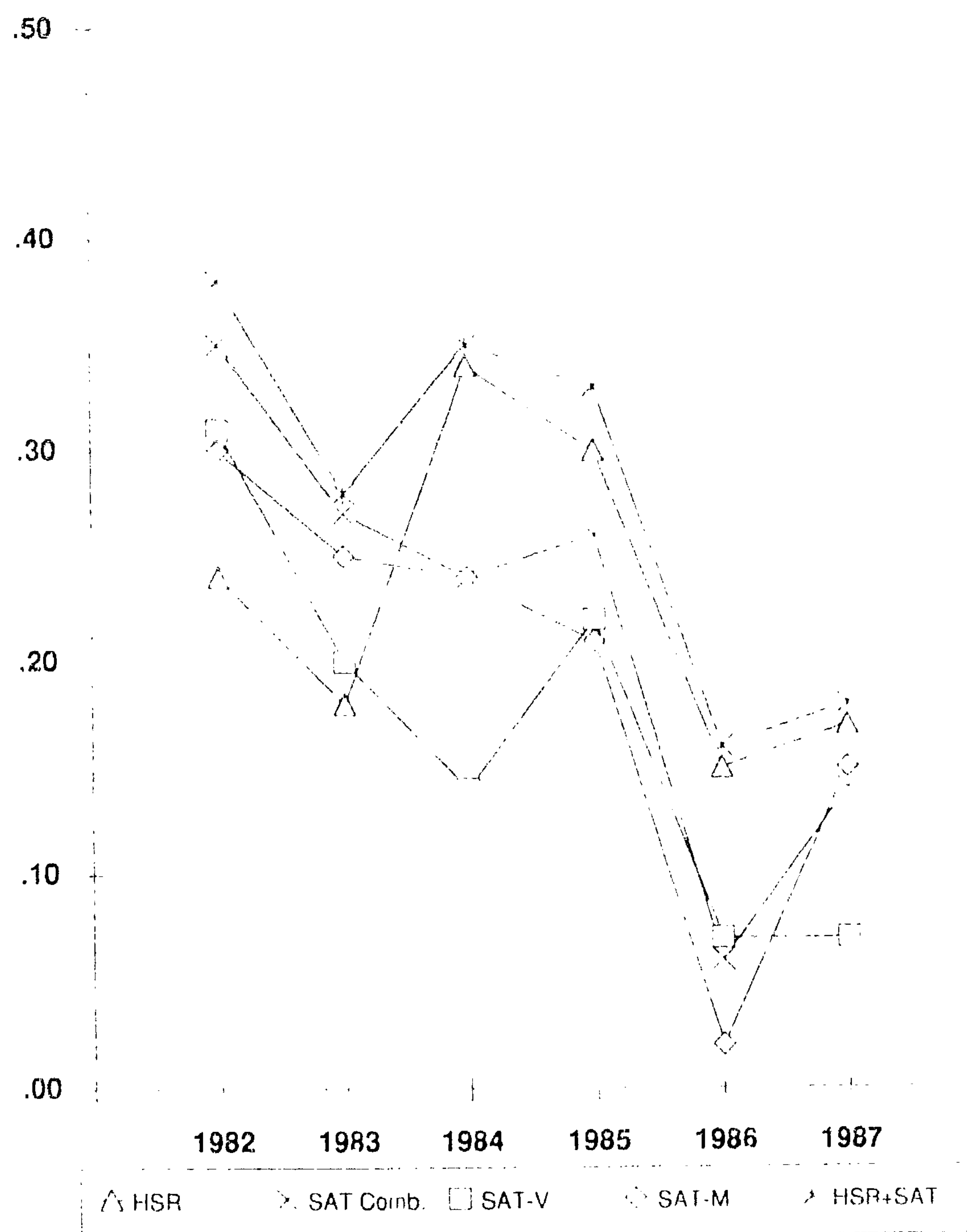


Figure CI-3A.

Correlations of SAT with Cumulative GPA for Men in an Eastern Private Research University

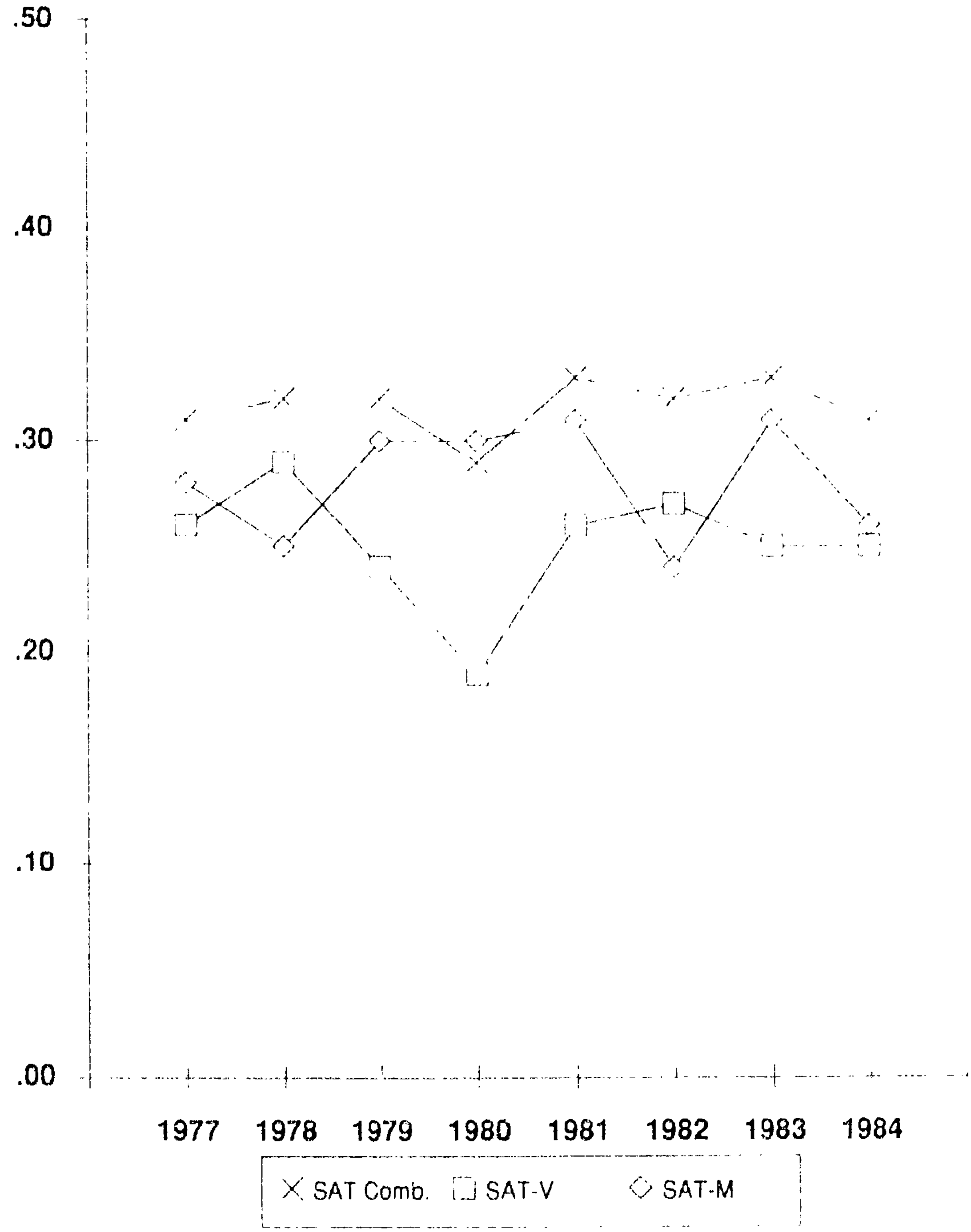


Figure CI-3B.

Correlations of SAT with Cumulative GPA for Women in an Eastern Private Research University

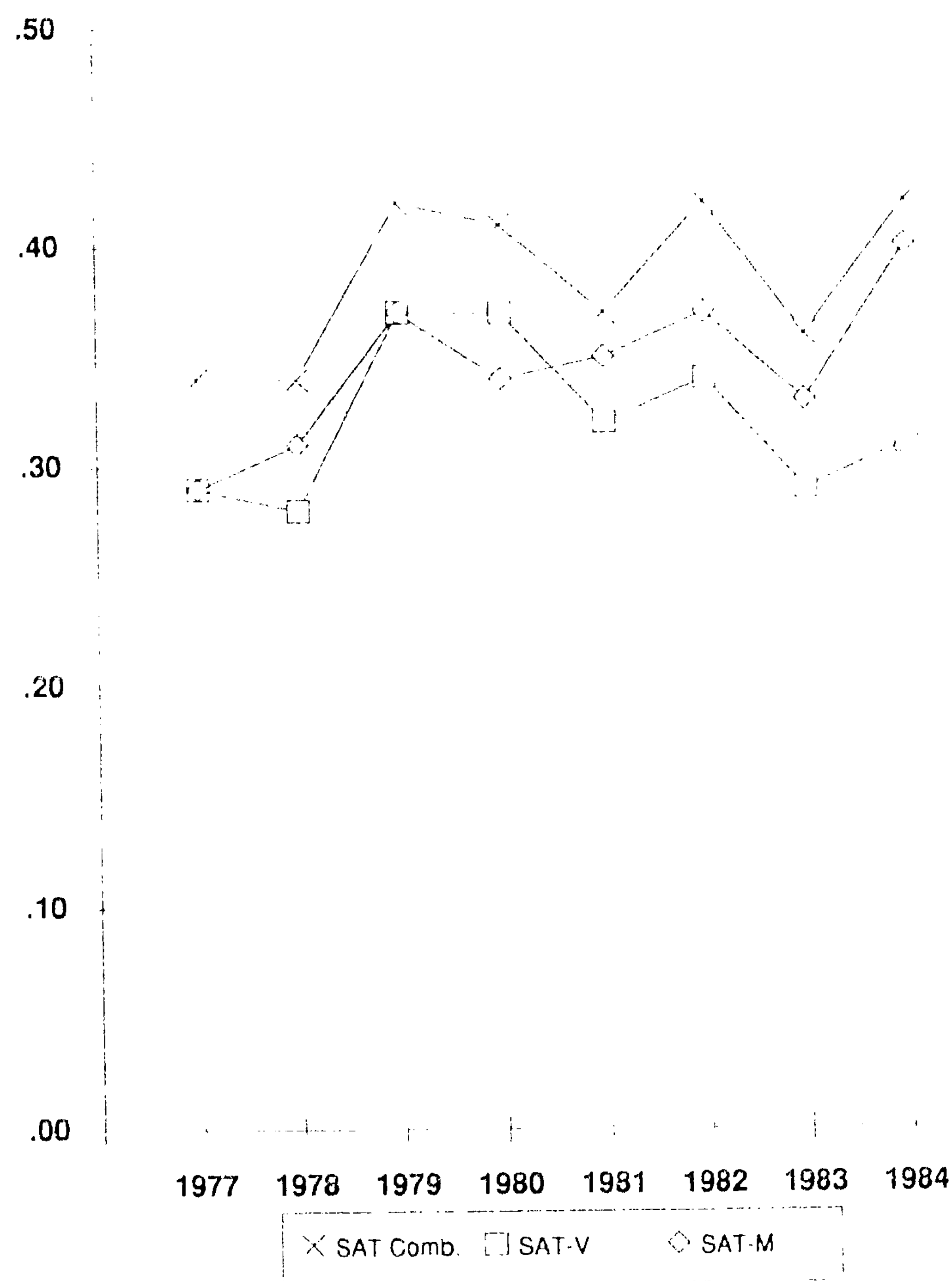


Figure CI-4A.

Correlations of HSR, SAT and HSR+SAT with FGPA for Men in a Midwestern Public Research University

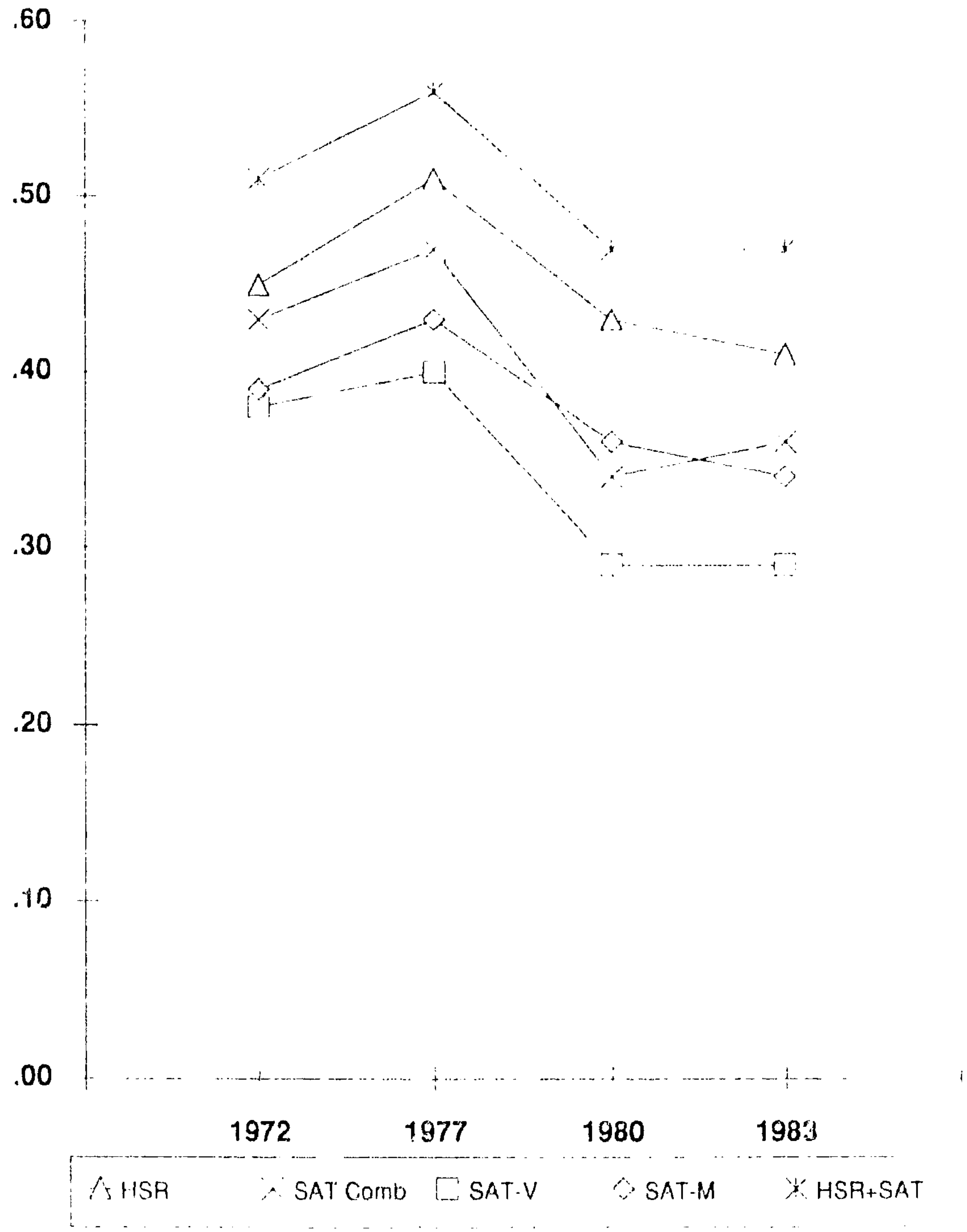


Figure CI-4B.

Correlations of HSR, SAT and HSR+SAT with FGPA
for Women in a Midwestern Public Research University

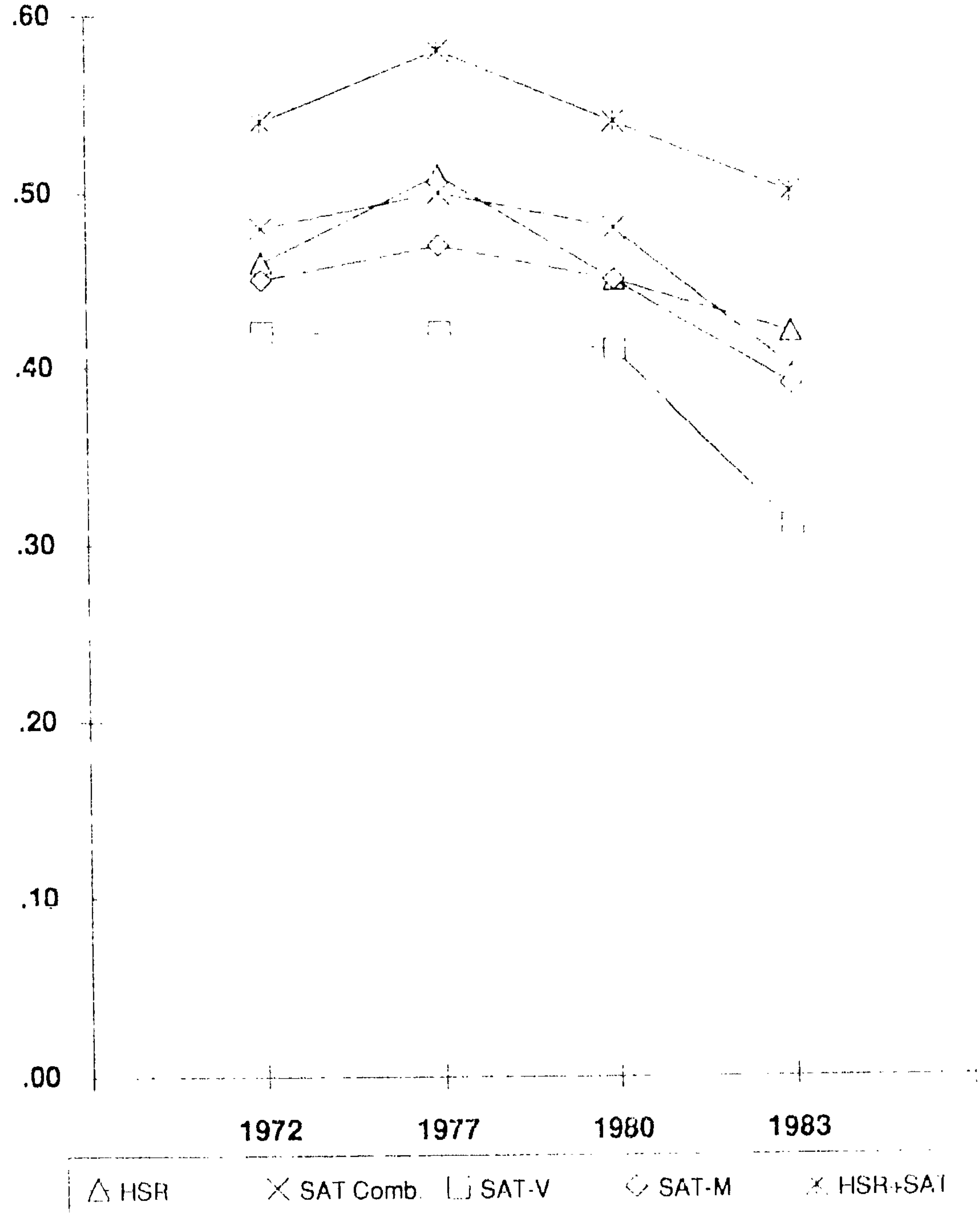
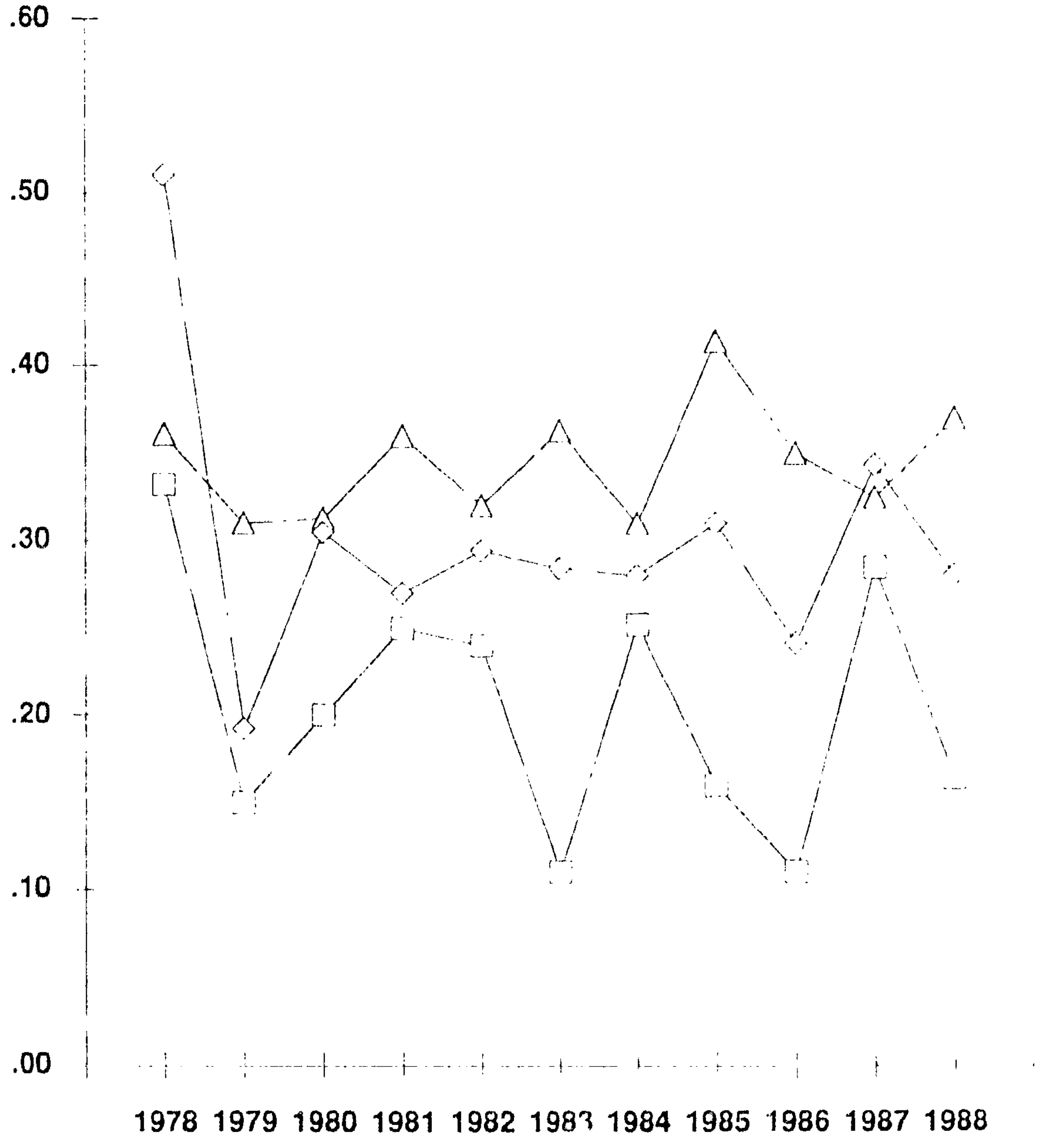


Figure CI-5A.

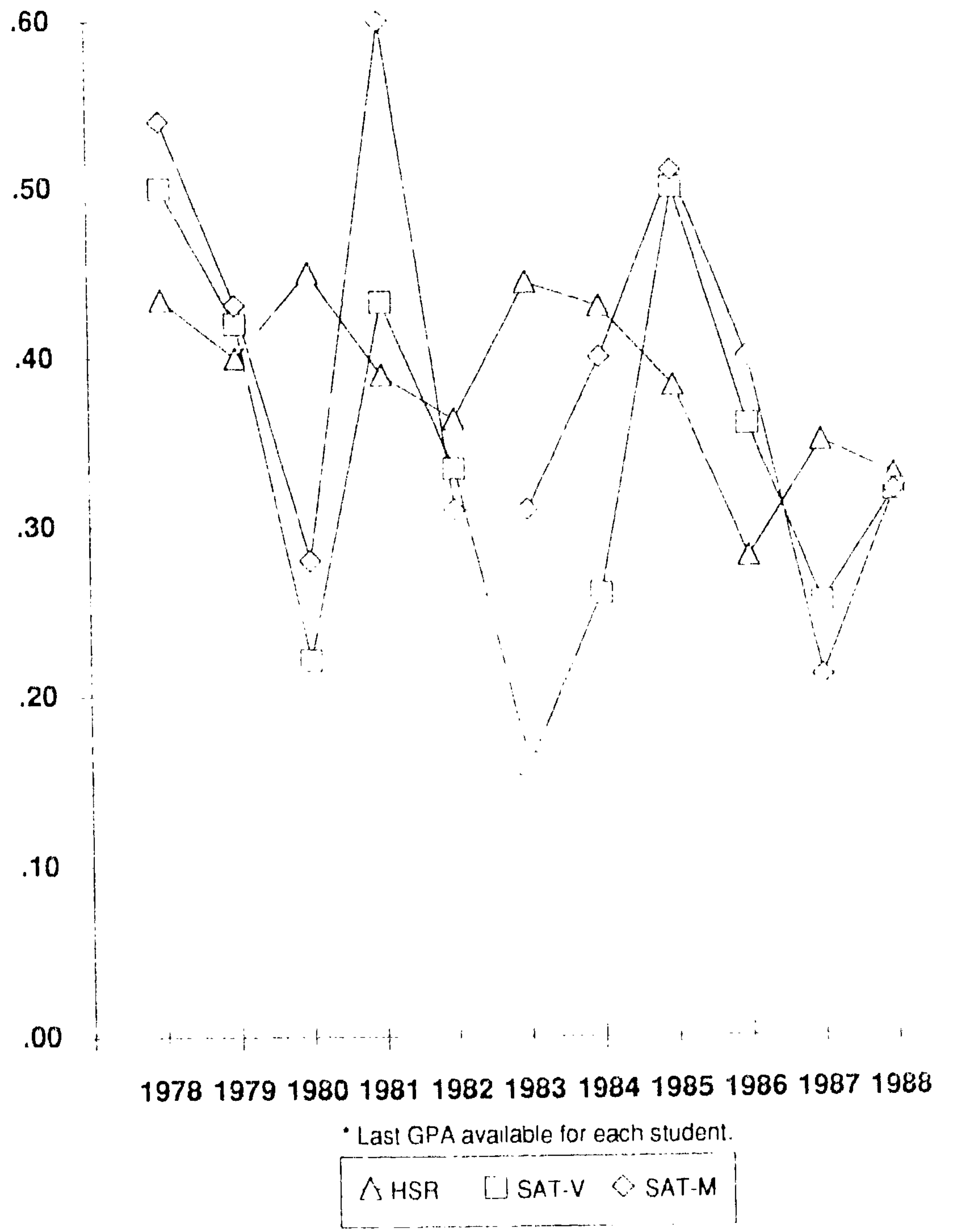
Correlations of HSR and SAT with GPA* for Men in a Midwestern Public Comprehensive University



* Last GPA available for each student.

Figure CI-5b.

Correlations of HSR and SAT with GPA* for Women in a Midwestern Public Comprehensive University



* Last GPA available for each student.

Figure CI-6A.

Correlations of HSA, SAT and HSA+SAT with FGPA
for Men in a Southern Public Comprehensive University

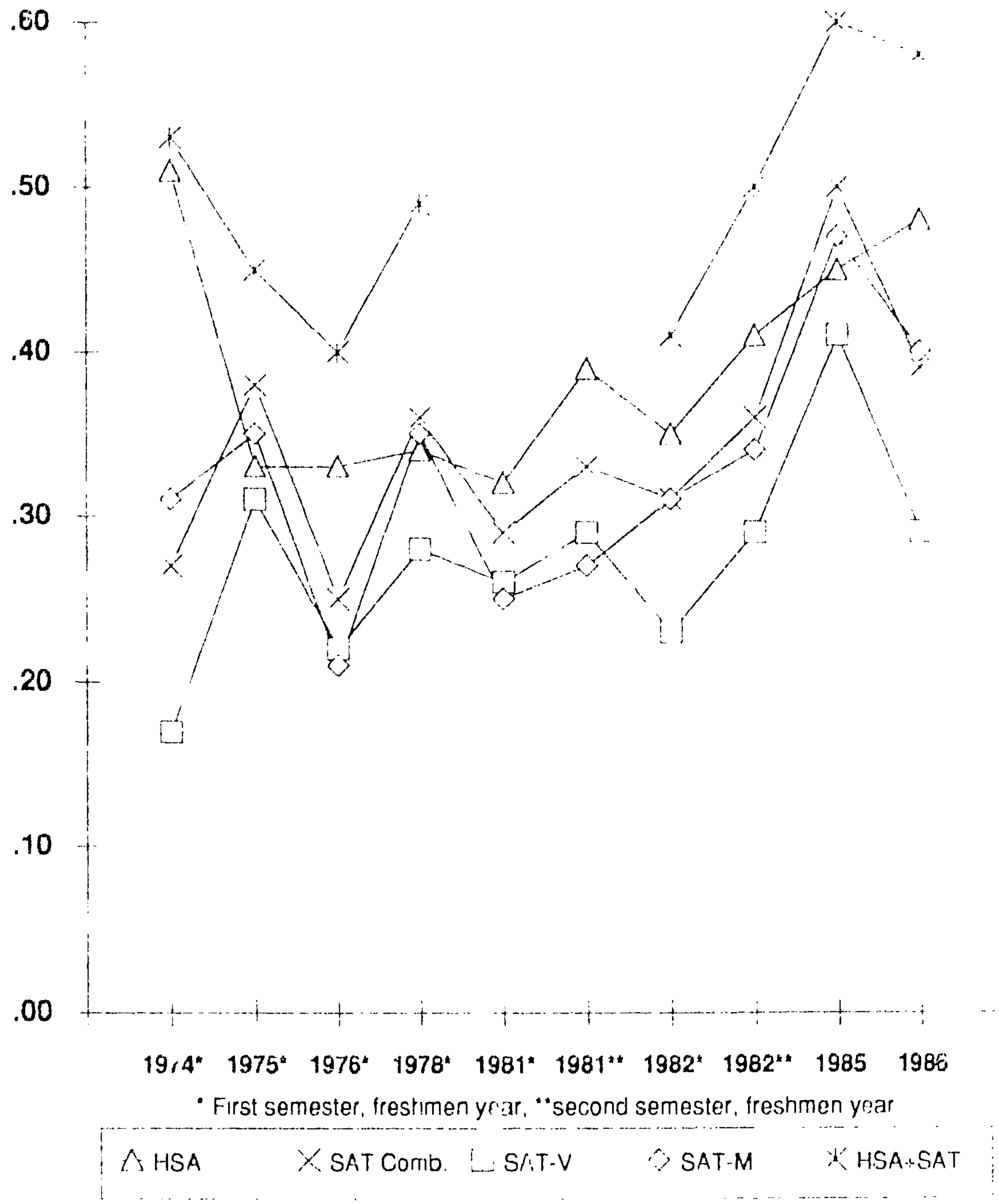
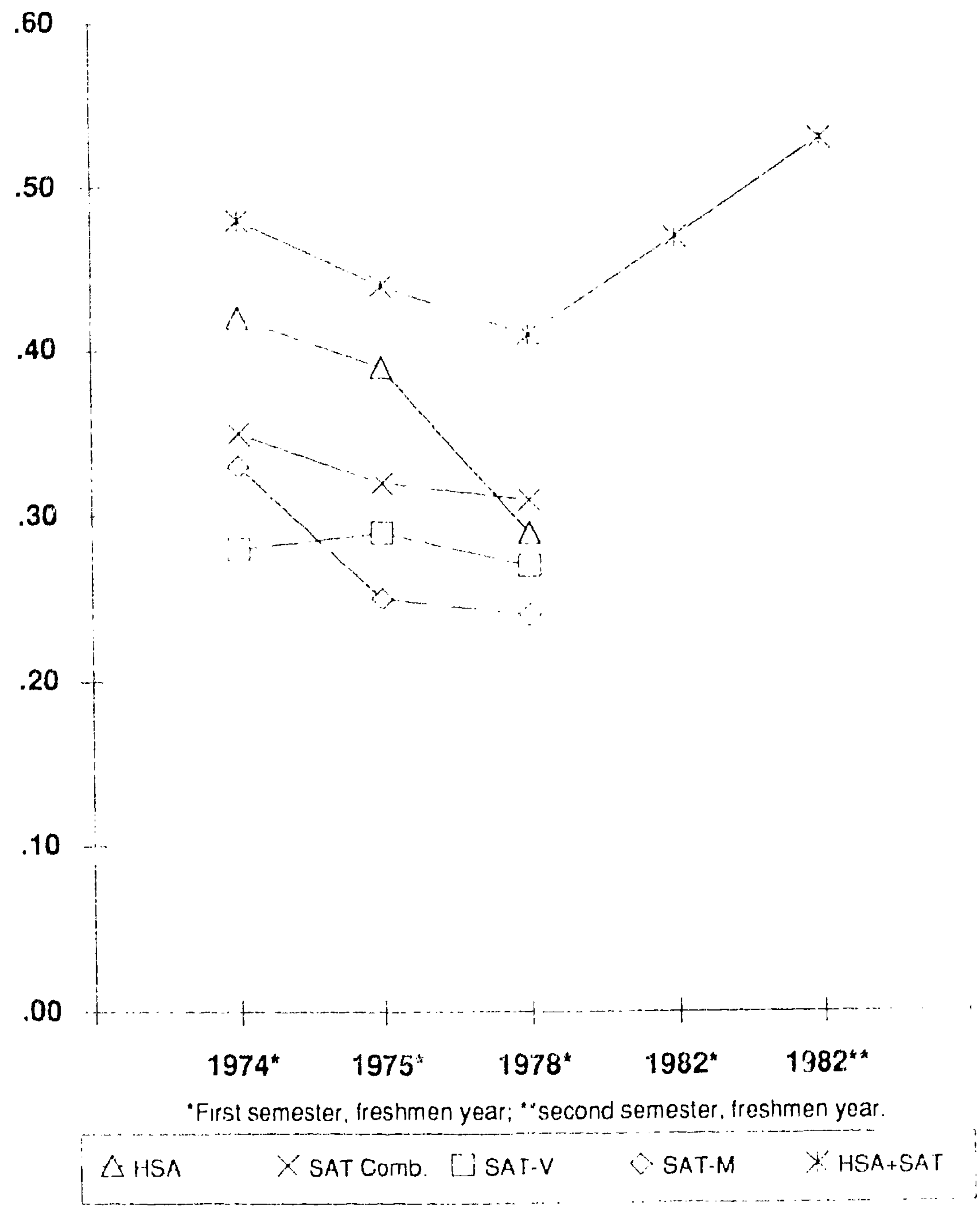


Figure CI-6B.

Correlations of HSA, SAT and HSA+SAT with FGPA
for Women in a Southern Public Comprehensive University



*First semester, freshmen year; **second semester, freshmen year.

Figure CI-7.

Correlations of HSA and SAT with FGPA
for an Eastern Public Research University

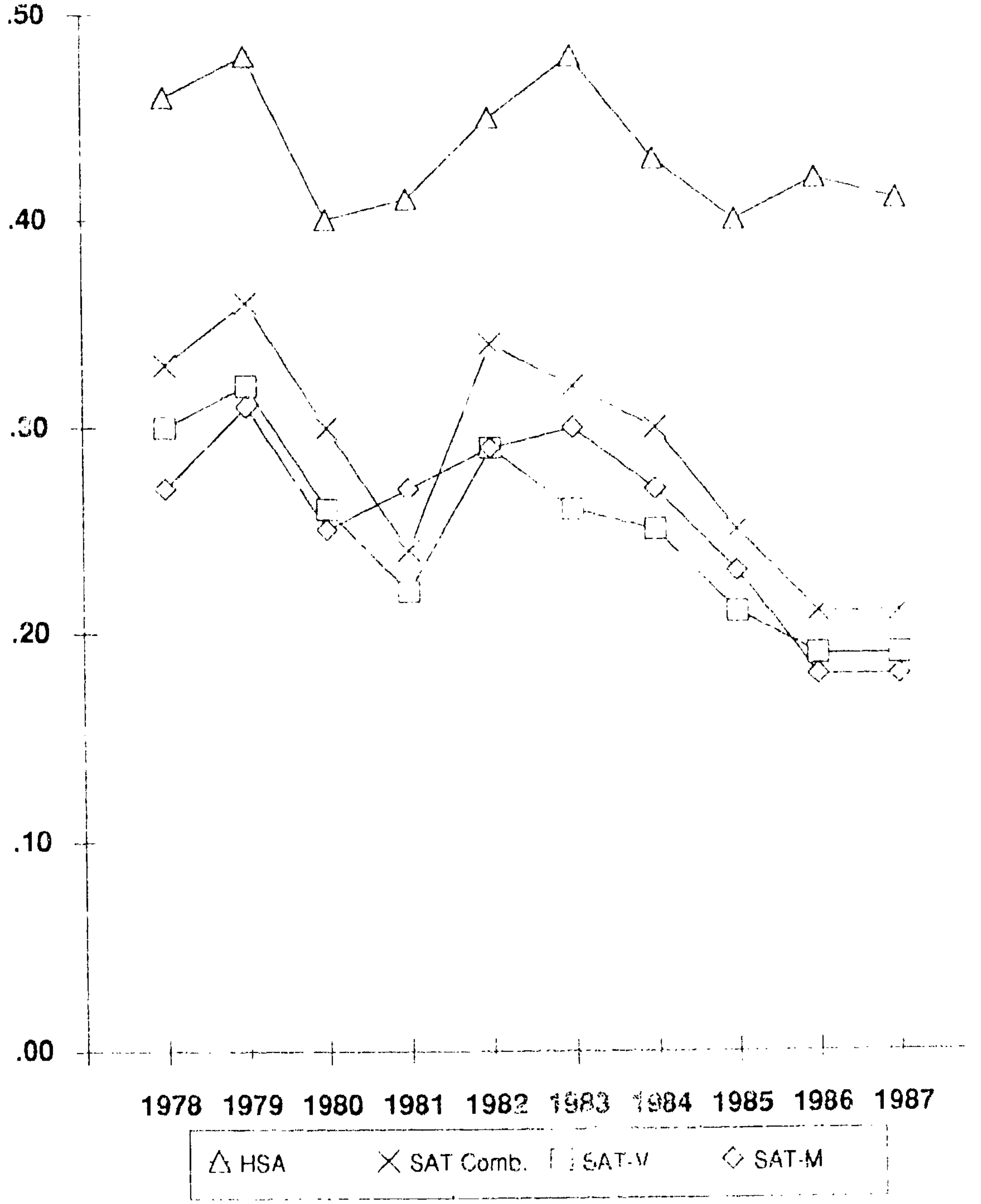


Figure CI-8A.

Correlations of HSR, SAT and HSR+SAT with FGPA
for Men in an Eastern Public Research University

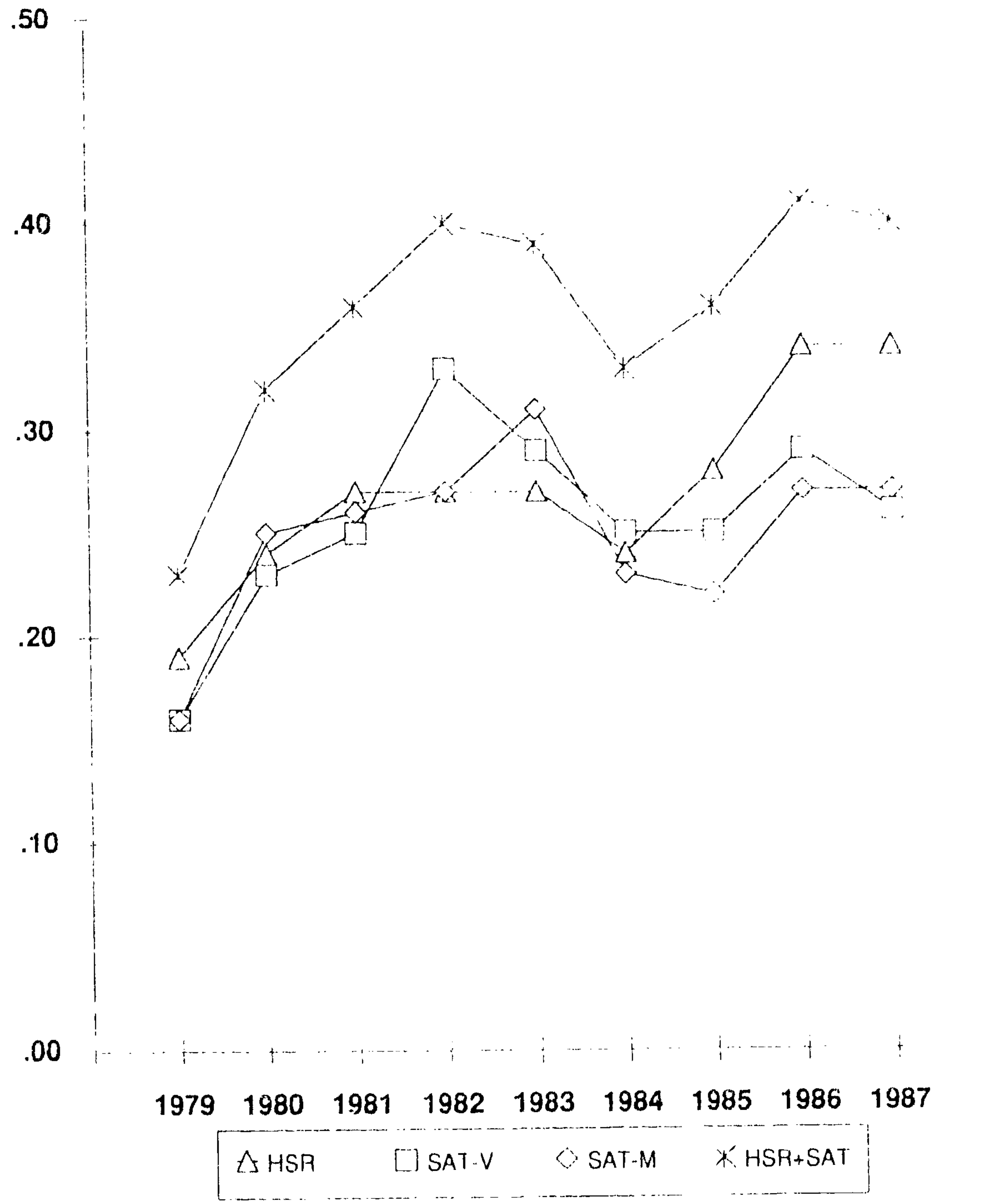


Figure CI-8B.

Correlations of HSR, SAT and HSR+SAT with FGPA
for Women in an Eastern Public Research University

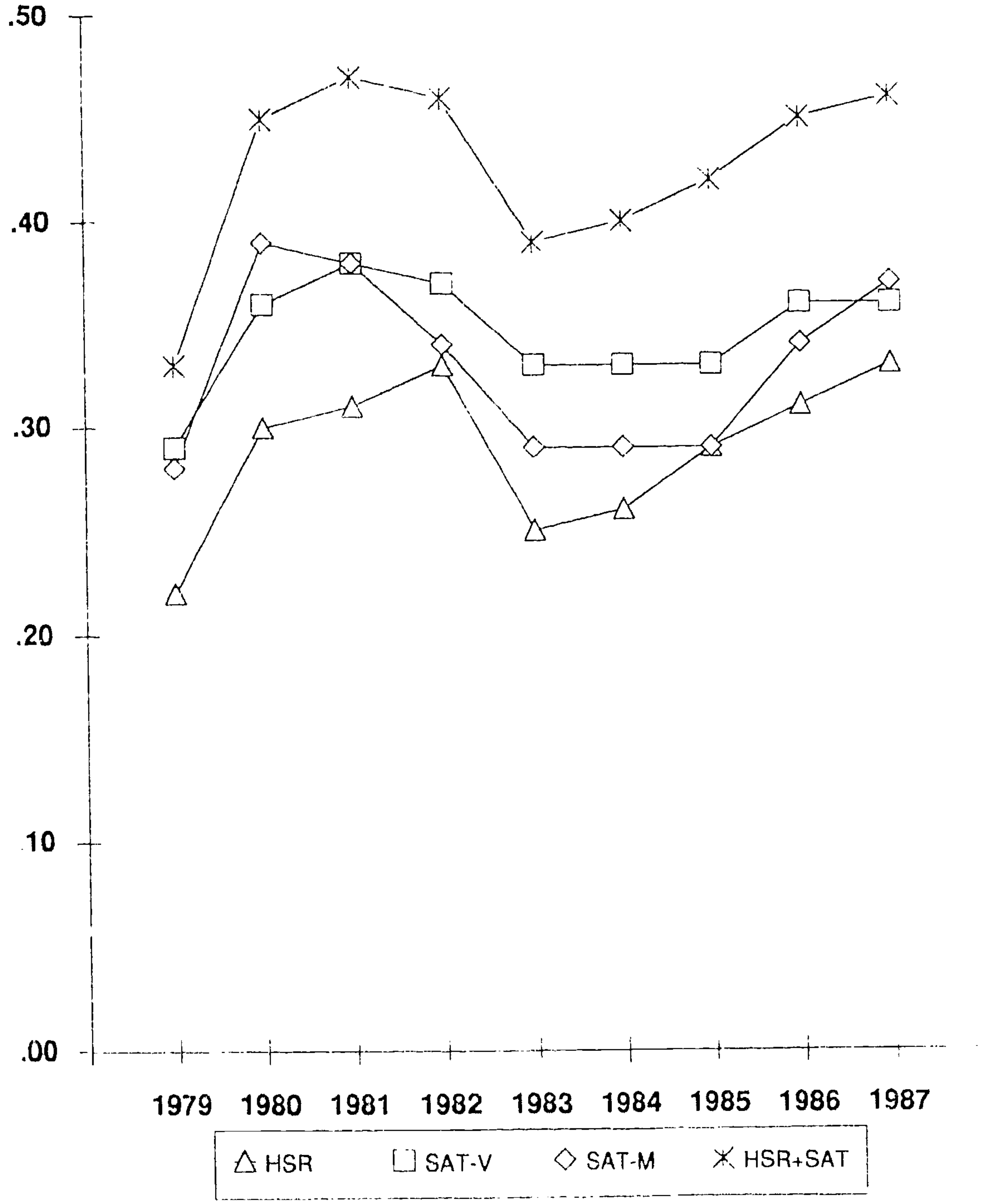


Figure CI-9.

Correlations of HSA and SAT with FGPA
in a Southern Private Comprehensive University

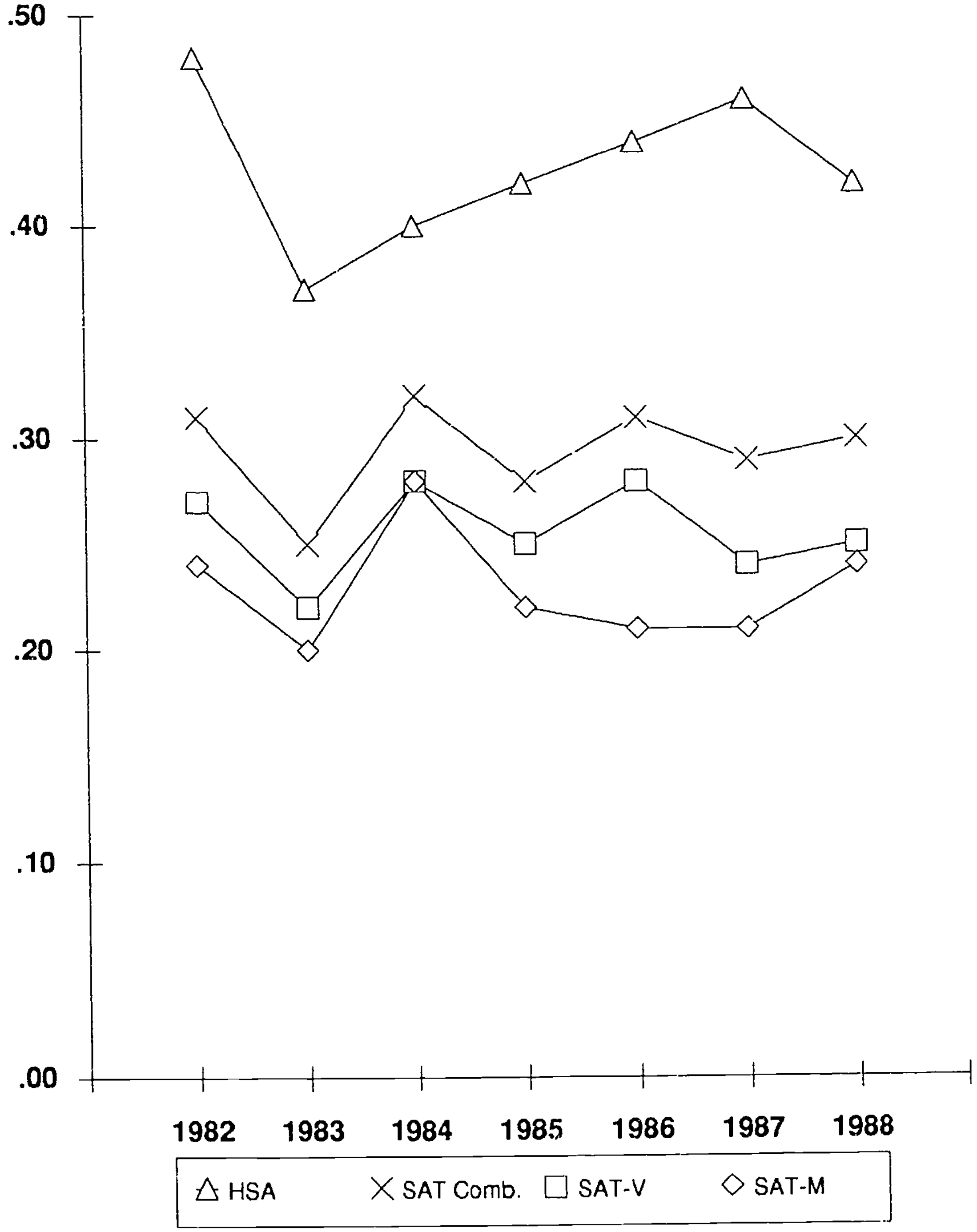


Figure CI-10A.

Correlations of HSA, SAT and HSA+SAT with FGPA for Men in a West Coast Public Research University

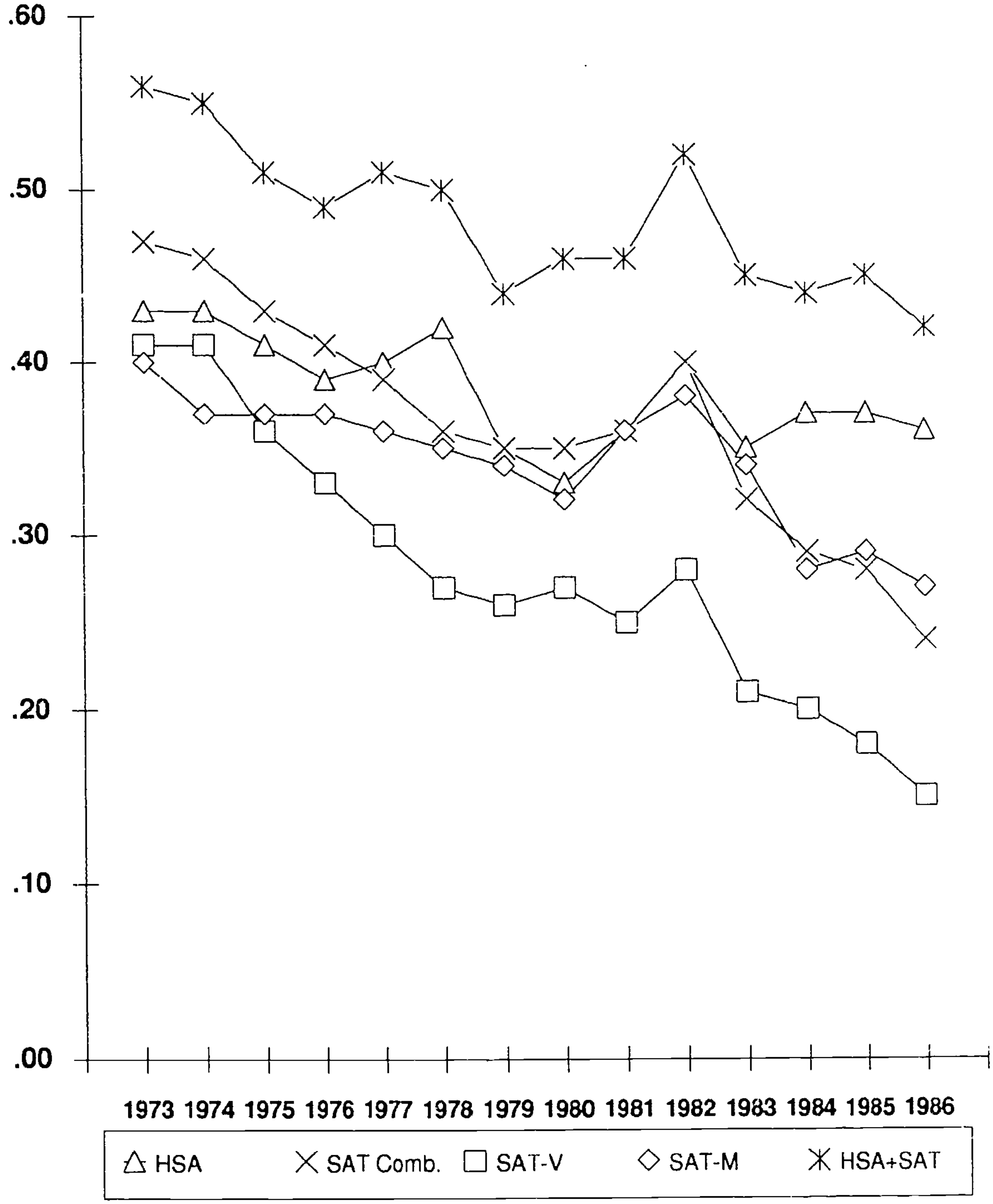


Figure CI-10B.

Correlations of HSA, SAT and HSA+SAT with FGPA
for Women in a West Coast Public Research University

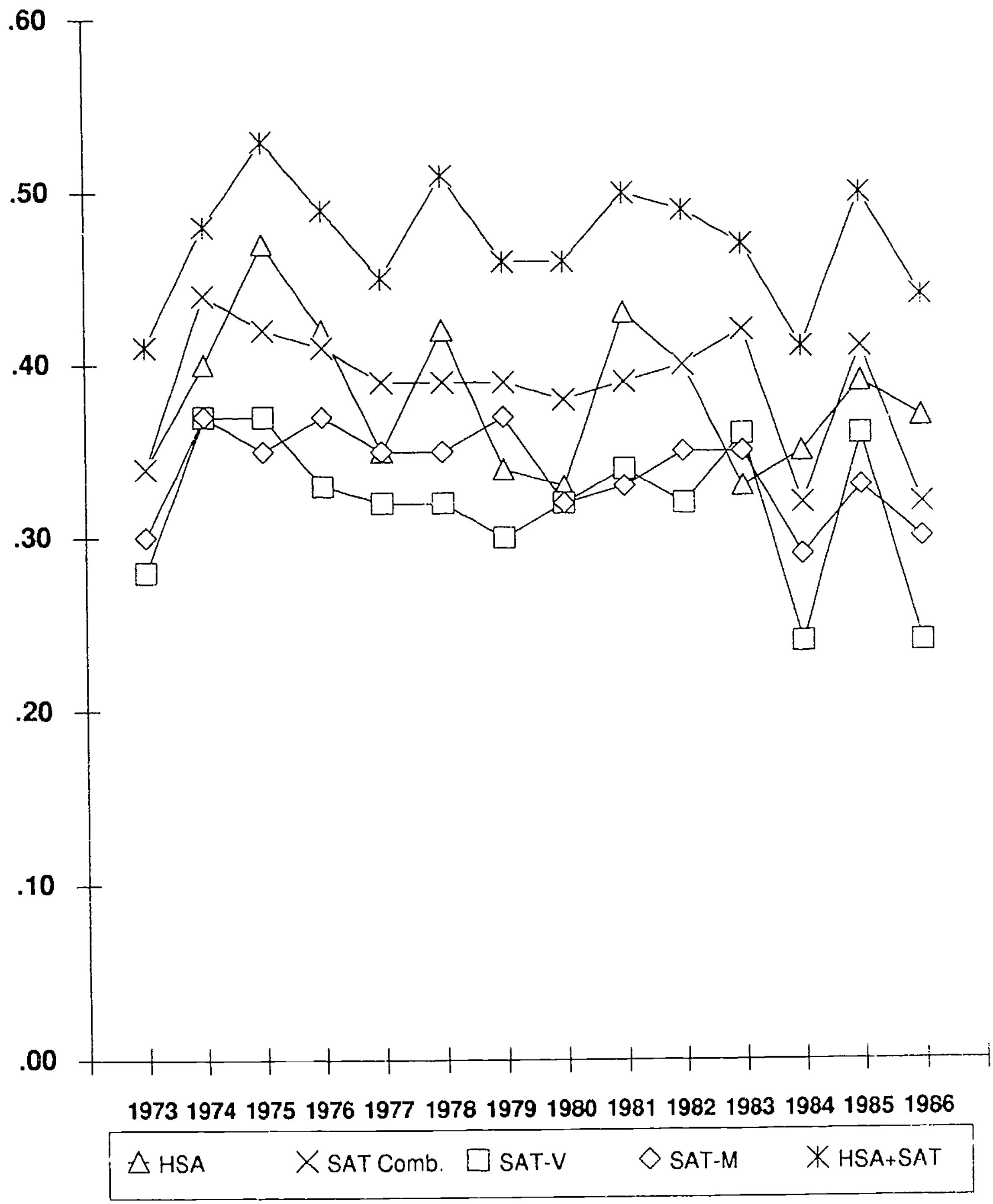
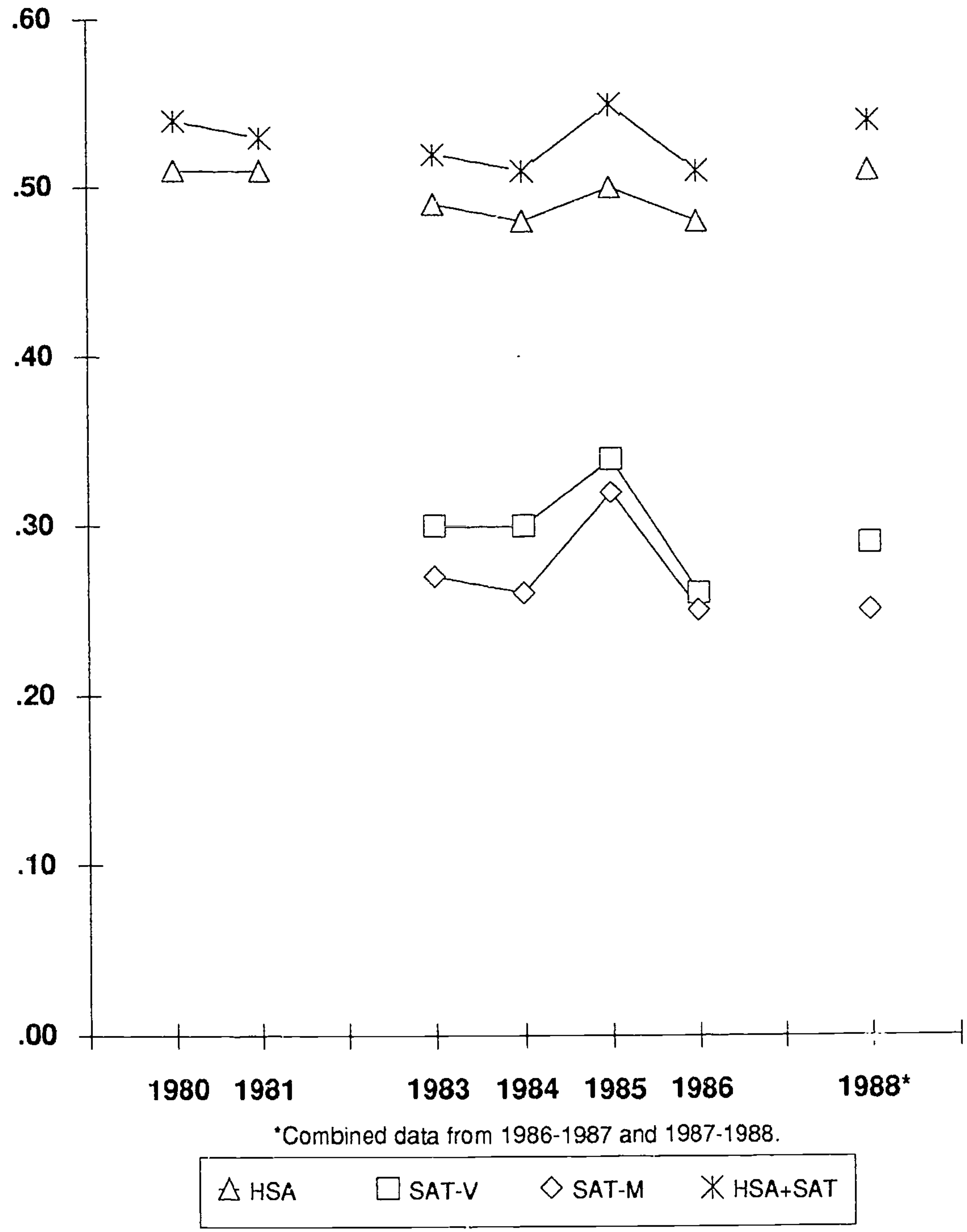


Figure CI-11.

Correlations of HSA, SAT and HSA+SAT with FGPA
in an Eastern Public Research University



*Combined data from 1986-1987 and 1987-1988.

Figure CI-12A.

Correlations of HSR, SAT and HSR+SAT with FGPA for Men in a Southern Public Research University

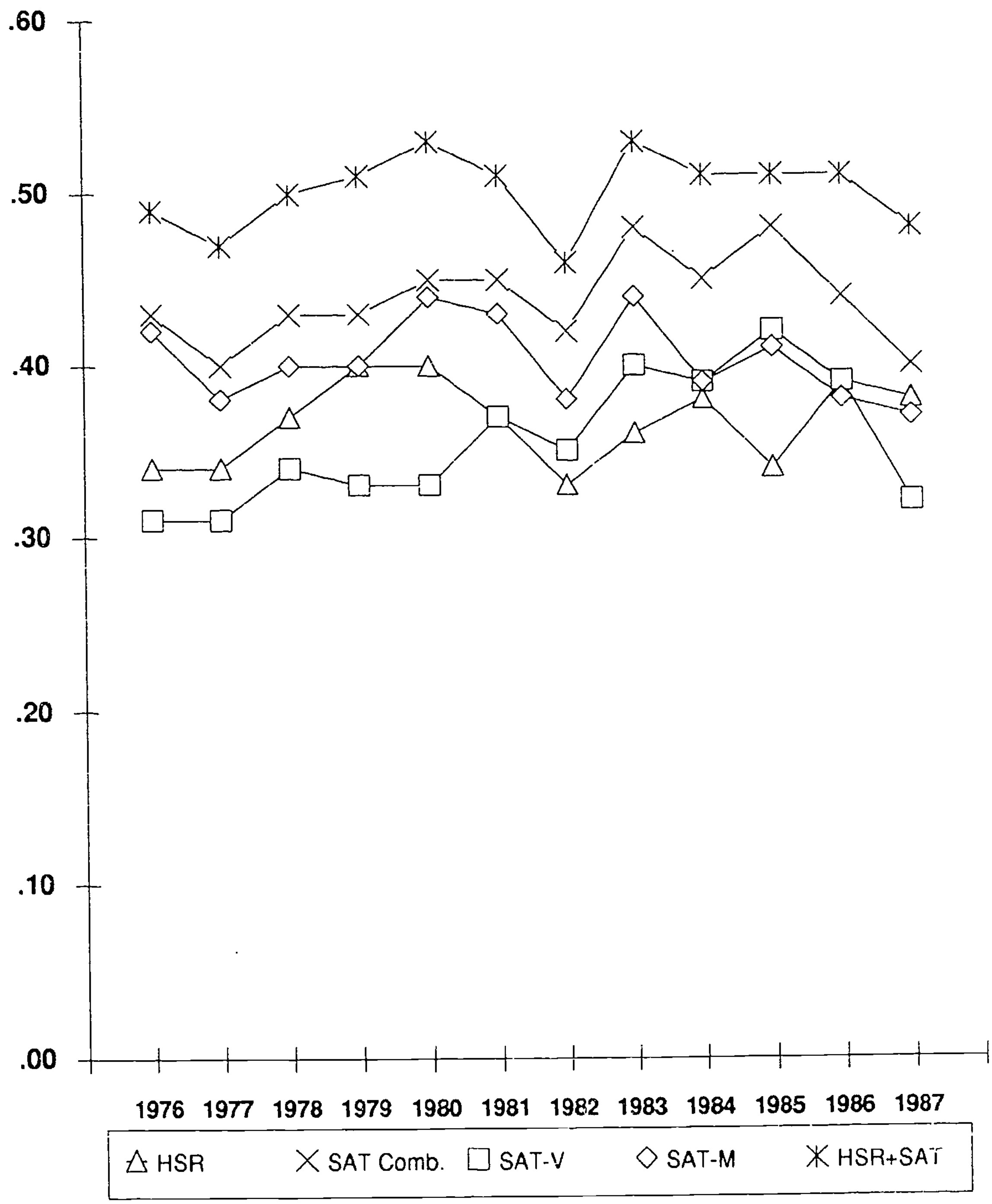


Figure CI-12B.

Correlations of HSR, SAT and HSR+SAT with FGPA
for Women in a Southern Public Research University

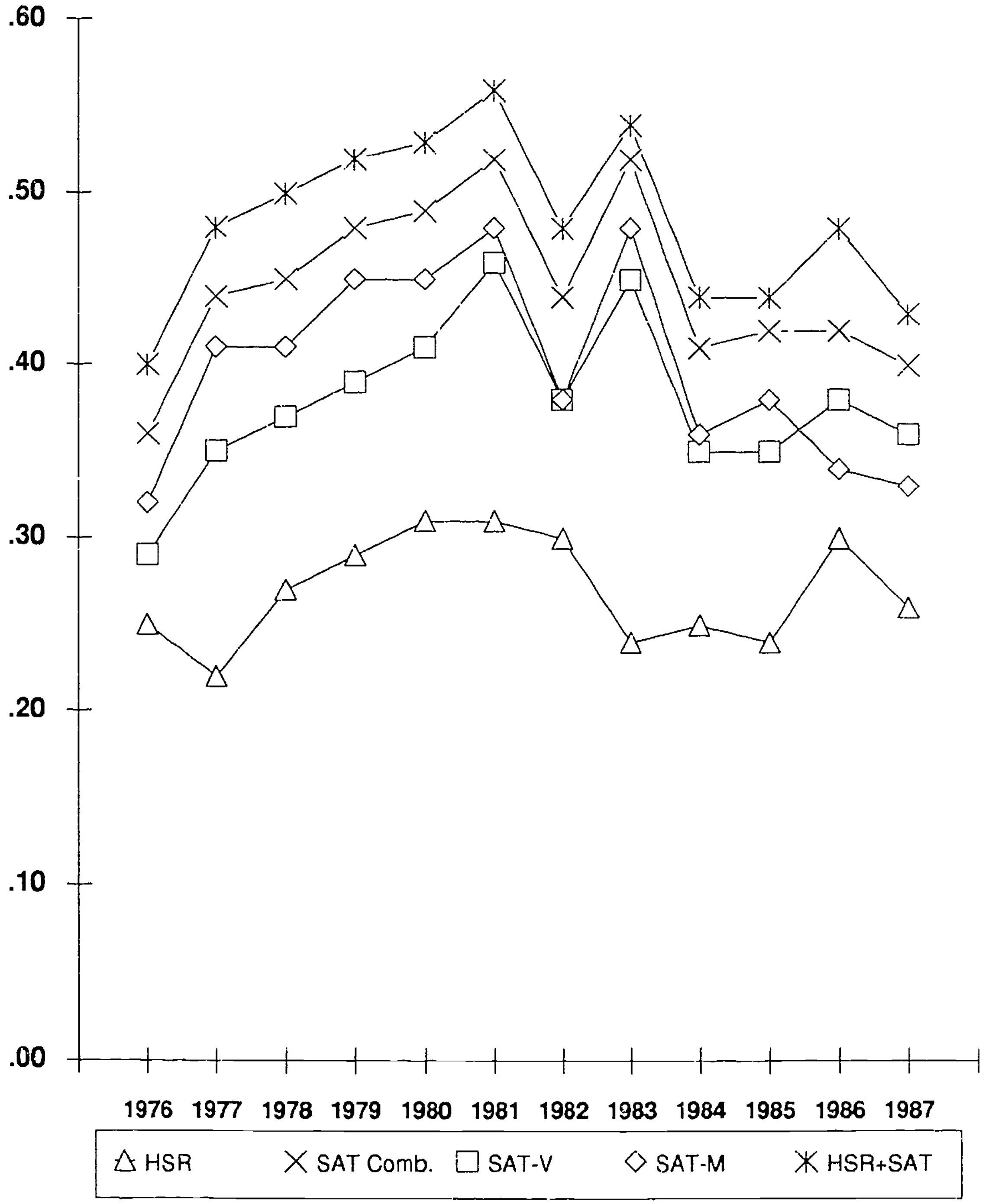


Figure CI-13.

Correlations of HSA, SAT and HSA+SAT with FGPA
in an Eastern Private Liberal Arts University

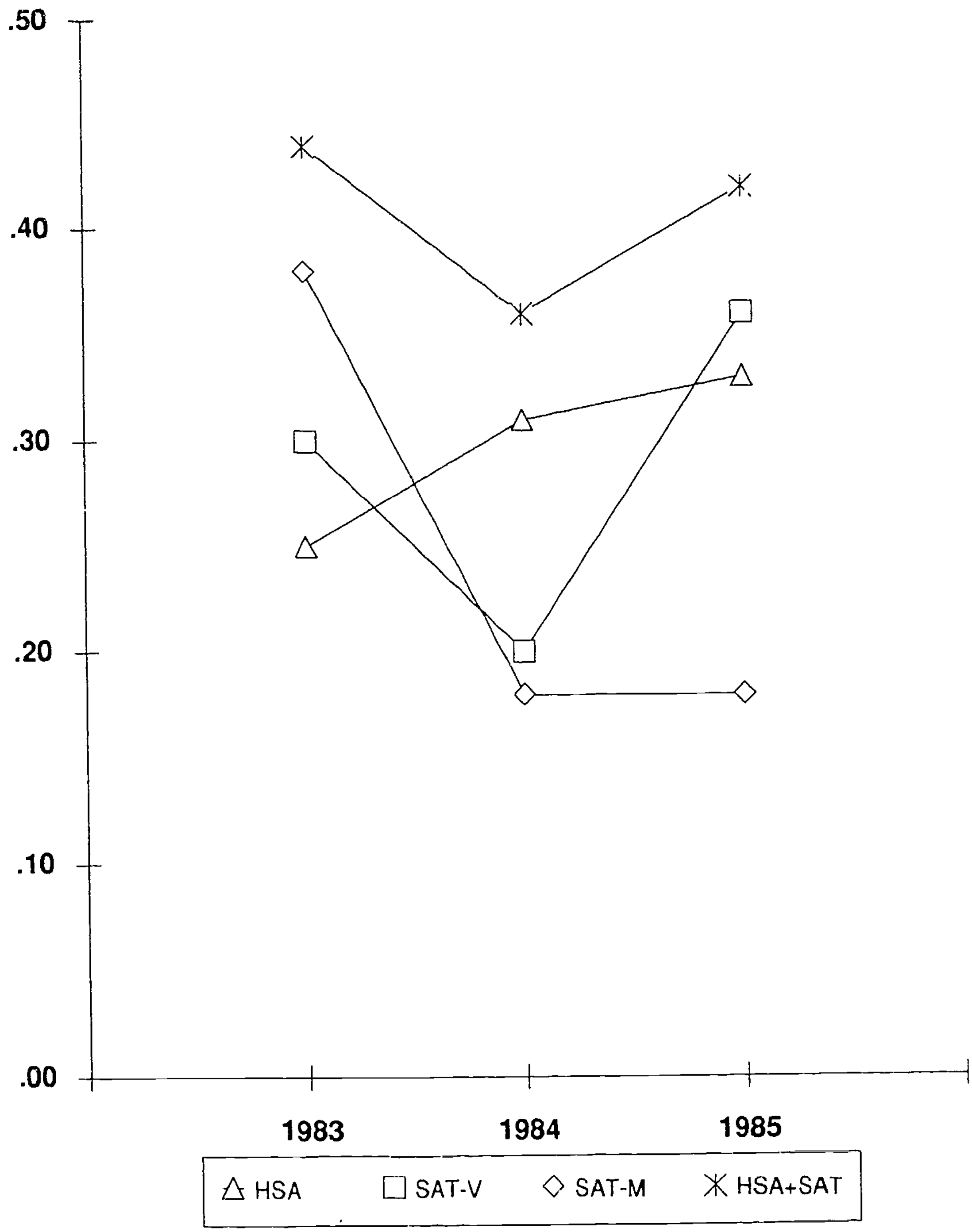


Figure CI-14A.

Correlations of HSR and SAT with FGPA for Men in a Southern Public Research University

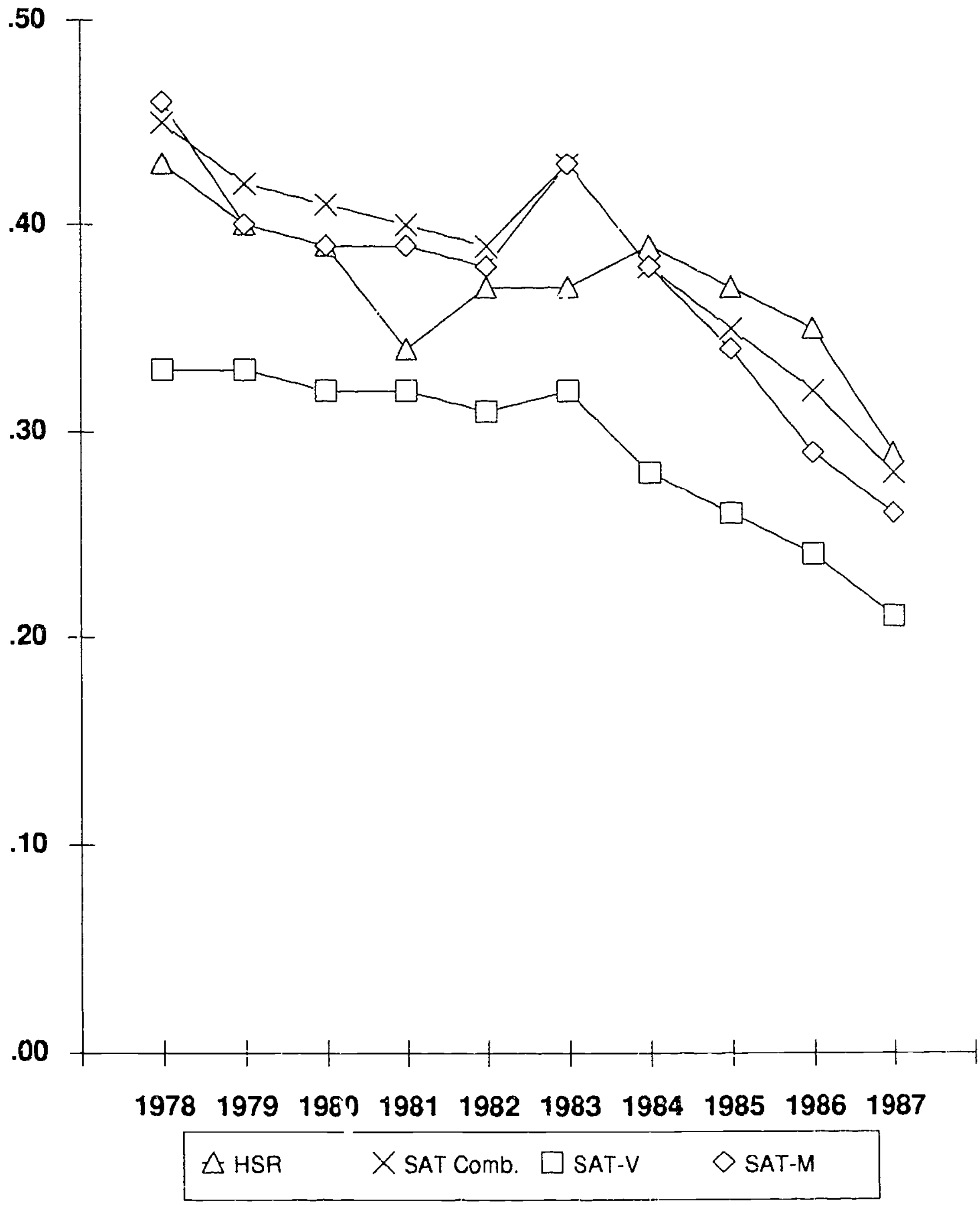


Figure CI-14B.

Correlations of HSR and SAT with FGPA for Women
in a Southern Public Research University