Going Beyond Traditional Examinations of the Relationship between High School Grades and College Performance

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Thomas Burke Kiger

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Nathan R. Kuncel, Adviser

October 2011

Acknowledgements

It is hard to believe that my graduate school career is coming to a close. It has been a great experience full of challenge and growth. Many people have aided me in this journey and deserve recognition.

First, I would like to thank my dissertation committee members: Drs. Nathan Kuncel, Paul Sackett, John Campell, and Mark Davison. I truly appreciate their help and insight during this process. Their input has greatly improved my work. In particular I would like to thank my adviser, Nathan, for his support and guidance over the last five years. His mentorship has helped me grow as a researcher and practitioner in the field of I/O psychology.

My friends have also been instrumental to my success. My graduate school experience was greatly enriched thanks to my cohort: Winny, Rena, Stacy, and Kara. We’ve certainly been through a lot together and I look forward to more adventures in the future. I would also like to thank the rest of my fellow graduate students over the years. Bowling with James, Adam, and Phil has been an endless source of amusement. Attending Gopher basketball and football games with Winny and Amanda was always a lot of fun. I also want to thank Dr. Doug Molitor, for not only being a great boss but also a mentor and friend.

Finally, I want to thank my family for their love and support. My parents, Nancy and Burke, instilled in me the importance of education at an early age and have been a constant source of love and encouragement. My brother, Jack and sister, Katy, have been great friends and supporters over the years. I also want to thank my grandparents, Howie and Clarie, and my Uncle Steve for their love, encouragement, and support.

Abstract

The overarching goal of this thesis is to conduct a more detailed examination of HSGPA as a predictor of college performance. Specifically, following a review of the pertinent literature, three issues were considered. First, variability in the predictive power of HSGPA across institution was assessed including the extent to which college/university characteristics can predict such variability. Results indicated that institutional characteristics can help explain variability in the relationship between HSGPA and FGPA. Such information can result in better estimates of validity at specific institutions. Second, the degree to which HSGPA displays non-linearity in the prediction of college grades across its range was examined. Predictive strength was found to decline for HSGPAs above 4.0. Finally, the impact of socioeconomic status on the predictive power of HSGPA was assessed. SES had little impact for all but the lowest levels of HSGPA. A weaker relationship between HSGPA and FGPA emerged for students who were both low in SES and HSGPA when compared to their higher SES counterparts.

Table of Contents

List of Tables

List of Figures

Introduction

High school grade point average (HSGPA) is acknowledged as one of the most important predictors in college admissions. However, in many respects, little is known about how HSGPA functions. In contrast, the application of standardized tests, such as the SAT (Atkinson & Geiser, 2009), has received considerable attention and debate. Many who oppose the use of standardized tests suggest emphasizing HSGPA as an alternative metric or fail to explicitly provide alternatives to standardized test scores at all. The nature of HSGPA as a predictor and its suitability are often simply assumed. The level of scrutiny applied to standardized tests is seldom applied to HSGPA alone although in some cases parallel work has been conducted with HSGPA and/or the combined effects of HSGPA and the SAT (Shen, Sackett et al., 2011; Sackett et al., 2011).

The overarching goal of this thesis is to conduct a more detailed examination of HSGPA as a predictor of college performance. Specifically, following a review of the pertinent literature, three issues will be considered. First, variability in the predictive power of HSGPA across institution will be assessed including the extent to which college/university characteristics can predict such variability. Second, the degree to which HSGPA displays non-linearity in the prediction of college grades across its range will be examined. Finally, the impact of socioeconomic status on the predictive power of HSGPA will be considered.

The importance of understanding such a key predictor of college performance can hardly be overstated. Recently, declines in the United States (U.S.) college graduation rate and the associated drop in world rankings has been cause for concern (Callan, 2008). This is particularly troublesome given findings that greater numbers of future jobs will demand higher skills and education (Bureau of Labor Statistics, 2011). Greater knowledge of how predictors of college performance operate would result in better classification of students to colleges. Knowledge of HSGPA in particular could lead to improvements in the structure of U.S. high schools. For example, improvements or changes to curriculum and/or grading schemes could promote academic success at college and guide students to appropriate venues to pursue their higher education.

It is crucial to study the properties of any predictor carefully but especially one for which such important ramifications exist. Such study is particularly germane due to the unique nature of HSGPA. While HSGPA clearly has a strong cognitive ability component, existing research has shown it is not simply a redundant measures of cognitive ability (i.e., standardized tests). Instead it appears to also capture effort over multiple years. Further, the maturational component of HSGPA also makes it unusual compared to other predictors that are collected at a single point in time. This component may cause unique issues due to key changes occurring for students during this time period of their life. High school students are typically faced with additional responsibilities and greater freedom as they transition into adulthood. For example, many students learn how to drive and have their first jobs while in high school.

HSGPA is considered by some to be an ideal predictor of college success since it is based on four years of work, measures both typical and maximal performance, and is a composition of evaluations from multiple raters. Further, it reflects academic activities that mirror those of future college work (i.e. writing papers, taking tests, and time management). This is even more applicable to college preparatory and Advanced Placement coursework. Employers would be delighted to have such a rich and varied analogue for making hiring decisions. In addition, high school grades might be preferable for the social meaning attached to their use. Geiser and Santelices (2007) note that achieving high grades over the course of four years of high school is based upon not only raw intelligence but also motivation, personal discipline, and perseverance. The use of HSGPA as a predictor promotes and rewards those who have demonstrated these qualities. Further, Geiser and Santelices note that grades are viewed as earned through effort while standardized tests are perceived to solely reflect raw intellectually ability. In addition, high schools grades are likely to partially reflect investment into subject matter. Catell's (1971; 1987) investment theory argues that there is both fluid intelligence (i.e. raw processing power) and crystalized intelligence (i.e. acquired knowledge and skills). Under this framework, the best tests of general mental ability are thought to be material an individual has spent time and effort learning or material that does not require specific knowledge. Some of the superiority of HSGPA as a predictor over standardized tests may lie in its closer testing of acquired knowledge.

The nature of both standardized admission tests and HSGPA has been well-established in many areas. Numerous studies report a strong relationship for standardized tests, HSGPA, and a composite of the two variables with college grades (Burton & Ramist, 2001; Sackett et al., 2011). After applying range restriction corrections, a composite of the SAT and HSGPA accounts for roughly 40% of the variance in college grades. This is an impressive figure given the host of other variables that may impact the grades a college student earns.

Because of these strengths, it is not surprising that HSGPA is widely used in educational and employment settings. HSGPA or class rank was rated as the most important factor in a survey of 957 four-year institutions conducted by ACT, the Association for Institutional Research, the College Board, the Educational Testing Service, and the National Association for College Admission Counseling Admissions test scores (Breland, Maxey, Gernand, Cumming, & Trapani, 2002). Similar surveys conducted in 1979, 1985, and 1992 found HSGPA to be the most important factor. This ranking is supported by the widespread finding that the HSGPA is the strongest predictor of college grades (Ramist, Lewis, & McCamley-Jenkins, 1994; Zwick & Schlemer, 2004).

Grades are important beyond educational arenas as most employers and hiring managers (approximately 70%) use grades to screen and evaluate applicants (National Association of Colleges and Employers, 2005). This focus in organizational settings is not unreasonable given the moderate but useful relationships between grades and job performance (Roth, BeVier, Switzer, & Schippmann, 1996).

However, HSGPA is not an ideal predictor. One of the chief concerns is a lack of standardization. High school coursework is not universal or even consistent among students within the same school. An identical grade point average can be attained by two students with effectively no overlap in courses. High school courses can differ in a number of important ways, such as academic rigor, classroom structure, and teaching styles. Great diversity also exists across high schools in the U.S. including substantial variance in the number and types of advanced coursework available. Some high schools offer separate college preparatory tracks of courses that address more advanced material and/or cover material in greater detail. Further, some high schools also offer Advanced Placement courses which aim to prepare students for completion of Advanced Placement tests. High scores on these tests typically allow students to test out of lower level college/university courses. These differences can impact how well HSGPA works as a predictor. Bassiri and Schulz (2003) found that HSGPAs from less challenging high schools showed a flatter relationship with college GPA than HSGPAs from more rigorous high schools.

Teachers may also be a source of variance through favoritism, discrimination, and/or indifference towards certain students. There is a long history of concern surrounding the study of idiosyncrasies and subjectivity in grading practices, dating back to Joseph Rice’s work examining grades for spelling and mathematics exercises (Hillegas, 1912). However, recent work by Beatty et al. (2011) showed that when aggregated, grades are quite reliable. They found that the college grades for two random individual courses is fairly low (mean intra-class correlation = .37). However, when stepped up using the Spearman Brown prophecy formula, the reliability of FGPA rose to .86. But, the National Center for Education Statistics (NCES, 1984) found that course grading standards varied across the country. The most rigorous standards were found in the Middle and Southern Atlantic States. Roughly the same number of A’s and combined D's and F's (21 vs. 20 percent) were assigned in these regions. In contrast, twice as many

A's as D's and F's were assigned in the Pacific (30 vs. 14 percent) and West North Central regions (26 vs. 13 percent) of the U.S. Further differences have been demonstrated by type of course. Visual and performing arts, and personal and social development courses had about 2.5 times the number of A's than mathematics courses while giving out about a third as many D’s and F’s.

The related issue of grade inflation in high school is another concern for the use of HSGPA as a predictor. Rising average grades could diminish the predictor power of HSGPA. If greater numbers of students reach the maximum, HSGPA cannot be used to distinguish them. This has been shown to be a legitimate concern. The Cooperative Institutional Research found that 42.9% of incoming college freshman reported earning A averages in high school in 2000 (Higher Education Research Institute, 2001). In addition, the College Board found that the mean self-reported HSGPA increased from 2.97 to 3.92 from 1973 to 2002 (Camara, Kimmel, Scheuneman, & Sawtell, 2003). McSpirit and Jones (1999) found evidence of grade inflation even when controlling for a number of factors such as changes in chosen majors, shifts to more lenient course withdrawal policies, and improvements in student aptitude.

It is important to note that grade inflation might not be a spurious effect if student abilities are rising as well. For example, changes may have occurred in relevant factors such as student ability, course-taking patterns, and/or teacher effectiveness. Further effects could also be spurious if they are due to a rescaling (e.g. B’s are now B+’s, A-’s are now A’s, etc.). Some have argued that grade inflation might be a positive trend. Scocca (1998) argues that the rise in grades may be due to an increased emphasis on mastery of material versus sorting of students. If this is the case the predictive power of HSGPA might be improving due to a smaller influence of one's classroom peer group in the determination of grades. However, most researchers view the rise in grades over time as a negative occurrence (Camara et al., 2003). Hobart (1997) views grade inflation as a result of lowered standards while Brookhart (1998) argues that grade inflation is due to factors such as increased parental and student pressure and reluctance to assign low grades. More recent work by the NCES (2009) indicates that the trend of grade inflation may be ending. They found that overall GPAs increased from 2.68 in 1990 to 3.00 in 2009 but did not increase from 2005 to 2009.

Perhaps the more widespread use of weighted grades, which accommodate for the potential to gain HSGPAs above 4.0, has helped ameliorate this issue. The use of such a grading scale is by no means universal and can further muddle comparisons across high schools. A 4.0 at a high school without weighted grades does not mean the same thing as a 4.0 at a school with the potential to achieve grades as high as 5.0. In addition, in the calculation of HSGPA some schools use all courses while others use solely academic courses (Bridgeman, McCamley-Jenkins, & Ervin, 2000). These issues point to the need for a greater emphasis on academic rigor and support attempts to standardize the system for determining HSGPAs across the country.

Some attempts have been made to address these issues. The University of California system requires applicants to submit high school grades in a uniform format consisting of solely academic core courses (University of California, 1999). Their system explicitly excludes courses such as physical education, typing, and driver education and provides a bonus point for approved honors courses, AP courses, and International Baccalaureate Higher Level courses. This procedure generally resulted in improved prediction compared to traditional self-report HSGPA.

Criteria for success in college

Freshman Grade Point Average (FGPA) is commonly used as the criterion for success in college. This choice has both benefits and drawbacks. FGPA has the advantage of measuring performance in a more uniform college curriculum than typical in later years and allows for easier collection of larger amounts of data less affected by transfers and dropouts. However, it fails to provide information on college performance in later years and on the eventual completion of college (Zwick & Sklar, 2005). Findings from Burton and Ramist (2001) and earlier work by Wilson (1983) suggest that first year GPA and cumulative GPA function similarly and that FGPA is a reasonable proxy for cumulative GPA.

Similar to HSGPA, the use of college grades as criteria also results in issues due to differences in course and college/university difficulty. Stumpf and Stanley (2002) note that:

> If the total freshman population of a college is considered, it is remarkable that SAT I and ACT scores and high school GPAs have any predictive value at all across the entire freshman class, because many students do not take the same courses as others. Some courses are graded leniently, some stringently. At the end of the freshman year, a history-of-art major at a given college might not have been in a single course that a physics or engineering major has taken. (p. 1043)

Further, more gifted students (with higher admission test scores and HSGPAs) are likely to select more strictly graded scientific and quantitative courses. This could cause the correlation between these predictors and FGPA to be attenuated (Elliott & Strenta, 1988; Goldman & Widawski, 1976; Ramist, Lewis, & McCamley-Jenkins, 1994). Zak, Benbow, and Stanley (1983) illustrated this issue with their finding that the SAT predicted the cumulative 4-year GPAs of chemistry-major graduates at Johns Hopkins University better than grades across the entire freshman class. This improvement in prediction was likely due to the much higher overlap in coursework for the chemistry majors.

Examining differences in grading standards is crucial. Ramist, Lewis, and McCamley (1990) found that large differences exist by course type. The most harshly graded courses differed by over one grade point from those graded the easiest. Specifically, the courses that were graded the easiest were physical education (actual grades were .78 points higher than predicted), arts classes (including studio art, music, and theater at .56 points higher than predicted), and education (.50 points higher). The most harshly graded courses were in the hard sciences. Biology courses were the most difficult (.35 points lower), followed by science and engineering (.24 points lower), and calculus (.24 points lower). Ramist and colleagues have looked at statistical methods to adjust college grades for range restriction, variations in grading standards, and criterion unreliability (Ramist et al., 1990; Ramist, Lewis, & McCamley-Jenkins, 1994). Looking at 45 institutions they found an unadjusted correlation of .48 for the SAT and HSGPA with FGPA. After their corrections for range restriction, grading standards, and criterion unreliability, the validity jumped to .76.

Other criteria for success in college have also been examined. Degree completion has been particularly well studied. However, in part due to its binary nature, degree completion has been more difficult to predict using typical predictors of college performance. With a sample of nine institutions, Willingham (1985) found that preadmissions measures correlated .30 with graduation. This value increased to .40 when institution was added to the model. But, when looking within a single institution, the correlation was much lower (r = .15). Work by Bowen and Bock (1998) showed similar results. With the addition of variables not available prior to admission, the prediction of graduation can be greatly improved. Kanarek (1989) found a correlation of .60 by adding seven additional predictors (persistence, FGPA, math ability, written expression ability, total basic skills English, parents' education, and spoken expression ability) to the basic test scores plus high school record equation. Persistence to the sophomore year and first-year GPA provided most of this boost accounting for 62% of the variance explained by the model. A few studies have also looked at the relationship between the SAT and HSGPA with acceptance to advanced degree programs (Ph.D., law, and medical). Willingham (1985) found a moderate correlation (r = .32). Using data from Bowen and Bock (1998), Burton and Ramist (2001) calculated a slightly higher correlation (r = .39).

Fairness and bias concerns

From an ethical and fairness standpoint, predictors should not be biased due to unrelated factors. Colleges/universities also commonly desire to have a diverse student body that varies by race/ethnicity and social class. In addition, changing demographics make ensuring that admissions procedures are fair for a diverse set of students even more important. Larger numbers of Asian American, Hispanic, and African American students are attending institutions of higher education in the United States (Burton & Ramist, 2001). The number of non-traditional students is on the rise and the number of home-schooled students is rapidly rising. Older students and international students also introduce a host of unique considerations. Admission tools which are able to fairly evaluate students from such diverse backgrounds are crucial. Other concerns revolve around legal issues in the admissions process. Litigation in multiple states (such as Texas, Missouri, Louisiana) has forbidden the use of affirmative action, declaring that admission should not be based upon race, ethnicity, and gender (Burton & Ramist). These rulings have made selection procedures that aid diversity goals more appealing.

In the examination of issues of fairness and bias it is important to first define these terms. Even though popular usage often equates these terms, important distinctions exist in the psychological definitions of fairness and bias. Bias is a technical term referring to systematic group differences in item responses, test scores, or other assessments unrelated to the trait being assessed (Guion, 1998). Fairness is a value judgment of the appropriateness of a method of test use or a pattern of test outcomes (Sackett & Wilk, 1994).

Issues of adverse impact are a serious concern in college admissions. Diversity is a valued and important goal in the American education system. Predictors which do not promote or even detract from this goal are problematic. The SAT has come under fire due to claims that it is a biased predictor. Comparable examinations of HSGPA and its impact on diversity goals have been less prevalent. If in fact, HSGPA is a better predictor in this regard, greater weight for it as a predictor could lead to greater diversity in colleges and universities.

Over the past three decades numerous research studies have examined the issue of differential prediction (or predictive bias) by race in both educational and organizational contexts (Houston & Novick, 1987; Ledvinka, 1979; Sackett & Wilk, 1994; Schmidt, 1988; Schmidt, Pearlman, & Hunter, 1980). The Cleary (1968) model has been widely applied to determine the existence of predictive bias in cognitive ability tests. This model considers a test biased if the criterion score from the common regression line is consistently too high or too low within subgroups. Most studies using the Cleary model have focused on the SAT; two key findings have consistently emerged. First, they have found infrequent evidence of substantial slope differences among ethnic groups. Second, when predictive bias is found to exist, it has come in the form of intercept differences favoring the majority group (larger intercept values for the majority group than the minority). This research has shown that ability tests do not underpredict performance of minority groups. Rather when regression lines show intercept differences, minority group performance is typically overpredicted (Bartlett et al., 1978; Jensen, 1980; Schmidt, 1988; Schmidt et al., 1980; Kuncel & Hezlett, 2007; Kuncel & Sackett, 2007). Comparably fewer studies have been performed looking at HSGPA. But, this parallel work has found, similar to standardized tests, that the performance of minority groups is typically over-predicted (Shen, Sackett et al., 2011).

It is important to note that when comparing subgroups care needs to be taken to ensure that a fair comparison is made. For example, different subgroups might disproportionately attend different types of schools, major in different disciplines, and/or take different courses. In a few cases such differences have been found to play a role. For example, men have been found to more often take harshly graded mathematics and science courses in college than women (Burton & Ramist, 2001).

As noted above, most of these types of studies have examined differences between African American students and White students. In addition to consistently showing overprediction, some studies have also shown that in a model with the SAT, HSGPA has a smaller influence for African American students than it does for the total group in the prediction of college grades (Burton, Morgan, Lewis, & Robertson, 1989; Ramist, 1984). The cause of this difference in prediction has yet to be examined sufficiently. Fewer studies have looked at Hispanic students. However, they typically have shown slight overprediction for the SAT and HSGPA when compared to the total group line as well (Shen, Kiger et al., 2011). Shen, Kiger et al. also looked at the impact of language proficiency among Hispanic students. Little difference in the overall pattern was found between students who spoke English best and those who were bilingual in English and another language. For both of these Hispanic language groups, slight overprediction existed. Shen, Kiger et al. (2011) also looked at the Asian subgroup. They found evidence of a disordinal interaction between traditional predictors (SAT, HSGPA, and their composite) and FGPA. Specifically, the grades of Asian students with higher predictor scores were over-predicted while those with lower scores are under-predicted as compared to the White or common regression line. However, this surprising finding was largely driven by language proficiency. The regression line for Asian students whose best language spoken was English was nearly identical to that of the White or common line. Asian students whose best language was something other than English and those who were Bilingual in English and another language showed a similar pattern to that of the total Asian group (i.e. a disordinal interaction). Ramist et al. (1994) and Gafni and Bronner (1998) found similar results showing that standardized tests underpredicted the performance of non-native English speakers.

When looking at differences by gender, as mentioned earlier, it is important to note that men are more likely to choose harshly graded majors and coursework. In this case, comparing the validity of HSGPA for men and women without accounting for differences in course-taking practices could give the appearance of bias against women when none in fact exists. Generally speaking when no adjustments are performed, college grades are typically underpredicted for women (Stanley et al., 1992; Ramist et al., 1994; Willingham & Cole, 1997). In other words, the actual grades of women are slightly higher than that predicted by a common regression line including both men and women while the actual grades of men are slightly lower. However, the extent to which such underprediction occurs is reduced when college grades are adjusted by course. Specifically, Ramist et al. found that the use of the SAT and HSGPA resulted in an underprediction of .06 for women on their FGPA. When course-taking patterns were accounted for the underprediction shrunk to .03. Further, when looking at different types of institutions, Ramist et al. found that for selective institutions there was an underprediction of .01, for average institutions it was .03, and for less selective institutions it was .05.

Other concerns revolve around the role of socio-economic status (SES) and its relationship to predictors of success in college. While SES is certainly related to performance in meaningful ways such as greater exposure to enriched learning environments and less pressure to seek part-time employment during high school. A predictor that derives all or a large part of its predictive power from SES would be undesirable. Controversy surrounding the impact of SES in part led to the University of California system's questioning of the continued use of the SAT. Those against the use of standardized tests claim that they are no more than wealth tests and that the inclusion of SES in predictive models greatly reduces their predictive power (Atkinson & Geiser, 2009). However, research has found that the predictive power of the SAT and HSGPA is not substantially reduced when SES is included in the model predicting college grades (Sackett et al., 2011). While SES does not greatly diminish the predictive power of the SAT and HSGPA it still might play an important role. For example, SES likely captures key characteristics such as educational opportunity, school quality, peer effects, and other social factors that are related to development of abilities associated with academic performance.

Finally, many of these factors are likely to interact. For example, Bridgeman, McCamley-Jenkins, and Ervin (2000) argue that "socioeconomic categories, such as the highest educational degree earned by either parent, interact with ethnic categories in a way that makes it difficult to attribute results to ethnic as opposed to socioeconomic categories (p. 8)." In addition, Shen, Kiger et al. (2011) found that language group operated differently for different racial/ethnic groups.

Study 1: Variability in the HSGPA-FGPA Relationship

Several large scale studies have demonstrated validity generalization for HSGPA across schools (Bridgeman et al., 2000; Ramist et al., 1994; Ramist & Weiss, 1990). However, variance is present even among these studies. After applying range restriction corrections, Bridgeman et al. found a correlation of .55 between HSGPA and FGPA while Ramist and Weiss produced a slightly higher correlation of .58. Ramist et al. found an even higher correlation of .69 albeit with more extensive corrections (range restriction, grading standards and criterion unreliability). Further these studies show variance across college/university. After correcting for range restriction and criterion unreliability, Ramist et al. showed that in selective schools HSGPA correlated with FGPA more highly (.62 vs. .53). In addition, differences existed by school size with HSGPA showing greater predictive power for larger schools than smaller schools (.59 vs. .54). While these values are certainly useful, they might not necessarily explain much for a particular school unless attention is paid to moderators and sources of variability across school. With a mean value and SD, a common strategy is to acknowledge that absent further information one cannot know where a given school falls in the validity distribution. The examination of validity at the institution level is essential in the analysis of variability in validity. Colleges and universities differ in many important ways such as selectivity, size, and campus environment. These factors could affect the validity of HSGPA and their impact on validity merits attention. The current study aims to examine the extent to which institutional characteristics relate to the validity of HSGPA in the prediction of FGPA after corrections for statistical artifacts.

A few studies have examined sources of variance in validity of college admission processes. As noted above Ramist et al. (1994) found that validity differed by school size and selectivity. Morgan (1989) also found differences due to these factors. He looked at changes over time in the predictive power of HSGPA and the SAT for college grades from 222 colleges that conducted 778 different validity studies over a 10-year period from 1976 to 1985. Looking at yearly averages of this data, the correlation between the SAT and FGPA declined slightly over this time period from .45 to .38. However, roughly half of this variability was due to range restriction. After multivariate range restriction corrections were applied for the SAT and HSGPA, the decline shrinks to.51 to.47. Further, the data suggest that some of this decline can be explained by school characteristics: private colleges, smaller colleges, and more selective colleges showed smaller changes in predictive power than public colleges, larger colleges, and less selective colleges. A consistent pattern did not emerge in the HSGPA data. However, variability in the average correlation between HSGPA and FGPA was demonstrated with multivariate range restriction corrected values ranging from.55 to.58. Analyses looking at school characteristics related to this variability were not conducted.

More recent work by Shen et al. (in press) also illustrates the importance of looking at validity differences by college/university. They found meaningful variability in the validity of the SAT across schools even when FGPA was corrected for range restriction and unreliability. The SAT had a mean corrected validity of r = .50 and an sd =.10 (range =.22 to.76). Higher validity was related to more selectivity, higher tuition, and a more homogenous campus life. Lower validity for the SAT was found in larger schools, schools that rely more upon non-traditional selection tools, and schools with a higher percentage of traditionally disadvantaged minority students. Looking at both the combined effects of HSPGA and the ACT with a 290 college sample, Munday (1970) found high variability in their uncorrected validity. Further, like Shen et al., he found that the variance was related to school characteristics: the standard deviation of ACT scores (12%), the percentage of students living on campus or under college supervision (10%), size of incoming class (9%), mean ACT scores (4%), percentage of full-time faculty (2%), percentage of female students (2%), amount of institutional control (private, state, vs. community-controlled institutions; 2%), and percentage of faculty with doctorates (1%). However, Munday did not analyze the ACT and HSGPA separately nor did he determine the extent to which variability was reduced with statistical corrections.

The current study aims to further examine variability in the HSGPA-FGPA relationship using a more current large sample dataset. The extent to which variability exists across college/university in the HSGPA-FGPA relationship will be examined first.

Research Question 1: Does meaningful variability in the HSGPA-FGPA exist across college/university?

The extent to which school characteristics can predict such variability will be examined if meaningful variability exists. Specifically, the work of Shen et al. (in press) on variability in the SAT-FGPA relationship will be applied and extended to the HSGPA-FGPA relationship.

Research Question 2: Do school characteristics relate to variability in the validity of the HSGPA-FGPA relationship?

Method

The present study replicates and extends Shen et al.'s (in press) work which used school characteristics to predict variability in the SAT-FGPA relationship. Because this study is based on the same dataset and addresses similar questions (albeit using HSGPA as a predictor), the analytical approach bears a strong similarity to that described in Shen et al. The same set of school characteristics developed in Shen et al. was also employed.

Sample

The College Board collected information on a host of variables including HSGPA and FGPA from 63,241 subjects across 110 institutions' 2006 entering cohort (see Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008, for prior research using this dataset). These were selected, based on a sampling plan, to be geographically diverse, to include large and small schools, to include public and private institutions, and to cover a broad range of school selectivity. The 110 schools included 63 private and 47 public institutions. Mean freshman class size was 1305.5 (sd = 1286.5), with a range from 106 to 6462. The mean of the mean high school GPAs among entering students across the 110 schools was 3.56 (sd = .24), with a range from 2.86 to 4.05 (grades above 4.0 reflect extra credit for Advanced Placement and honors courses). A subset of this sample that included school reported high school grades was utilized reducing the sample to 49 institutions and total sample size to 63,241.

Measures

High School Grades (HSGPA). Self-reported HSGPA was provided from student questionnaires completed when taking the SAT. Students were asked to report their cumulative HSGPA in one of the following categories: A+, A, A-, B+, B, B-, C+, C, C-, D+, D, and F/E. Letter grades were converted to their corresponding numerical equivalent: A+ (4.3), A (4.0), A- (3.7), B+ (3.3), B (3.0), B- (2.7), C+ (2.3), C (2.0), C- (1.7), D+ (1.3), D (1.0), and F/E (0.0).

SAT Scores. College Board records provided scores on each of the three SAT subtests (math, critical reading, and writing) for freshmen students in each of the 110 institutions that make up the current dataset. The subtests were combined into a single unit-weighted composite.

Socioeconomic Status (SES). Student's also reported information regarding three SES variables while completing questionnaires when they took the SAT: father's education, mother's education, and family income. Father's and mother's education was split into the following categories: grade school, some high school, high school diploma, business school, some college, associate's degree, bachelor's degree, some graduate school, and graduate degree. These options were recoded to represent the number of years of education: grade school was coded as 8, some high school as 10, high school diploma as 12, business school or some college as 13, associate degree as 14, bachelor degree as 16, some graduate school as 17, and graduate degree as 18. Family income consisted of thirteen response options: less than 10,000, 10,000-15,000, 15,000-20,000, 20,000-25,000, 25,000-30,000, 30,000-35,000, 35,000-40,000, 40,000-50,000, 50,000-

60,000, 60,000-70,000, 70,000-80,000, 80,000-100,000, and 100,000 and greater. These responses were numerically coded using the midpoint for each income range. Next, the natural log of this value was taken to normalize the distribution. In the entire population of SAT takers father's and mother's education correlated .60. The relationship between family income and education was slightly lower: father's education (r =.42) and mother's education (r =.39). An equally weighted composite of the three SES variables was created to represent an overall SES score. Only individuals for which all three variables were available were used.

Freshmen GPA (FGPA). FGPA scores were provided by each school from official records. Freshman grades for nearly all schools reach a maximum of 4.0. One college had a plus minus system that permitted a GPA above a 4.0 although few students ever exceed that level.

Institutional characteristics. The Institutional characteristics developed in Shen et al. (in press) were utilized in the current study as well. Their collection method is described in the following:

Institutional characteristics were collected from the dataset described above and College Board (2008). College Board (2008) is an annual publication which provides information on a wide range of student, structural and admissions characteristics for public and private colleges and universities in the United States. Due to the large volume of available characteristics, they were trimmed to a subset of 34 based upon past research and conceptual links between characteristics and validity. To avoid issues with multicollinearity, composites of strongly-related variables were created. Initially, a factor analysis on all available institutional variables was run. However, like Talento-Miller and Rudner (2005), the results were not readily interpretable.

Therefore, to create interpretable factors an iterative rational-empirical procedure was used. First, factors were separated theoretically and thematically. Next, the strength of correlations for variables within each factor was examined as well as the extent to which they hung together in factor analytic results. Variables that did not display these characteristics were dropped from the analysis. For example, two themes of interest were the use of classic predictors (i.e., high school grades and class rank, standardized test scores) and the use of alternative predictors (i.e., recommendations, extracurricular activities, and essays). The use of interviews did not fit with either of these two factors so it was cut from further analyses.

This process resulted in seven total factors: classic predictors, alternative predictors, selectivity, size and financial need, cost, homogeneous campus life, and math and science preparation. The classic predictors factor relates to how much schools rely on traditional selection tools, such as high school records and test scores. The alternative predictors factor refers to how heavily schools rely on other selection tools, such as letters of recommendations and extracurricular activities. The selectivity factor represents how selective a school is in the admission process; this factor includes variables such as higher mean SAT and GPA, lower percentage of students admitted, and higher percentage of students who were ranked in the top ten percentile of their high school class. The size and financial need factor consists of the size of school's student body and the financial need of the student population. It consists of variables such as freshmen class size, number of full-time faculty, and number of students that received financial aid. The cost factor represents variables related to the cost of attendance at a university including in-state and out-of-state tuition and average cost of room and board. The homogeneous campus life factor captures the level of similarity in students' college life experiences consisting of variables such as percentage of students living on campus, part-time students, and out of state students. The math and science preparation factor represents college/university expectations regarding the number of math and science units students will have completed prior to applying for admission. Table 1 lists the institutional characteristics that make up each factor.

Each of the seven factors had fairly high internal consistency reliabilities: classic predictors (mean inter-item r = .74; 2 characteristics, α = .85), alternative predictors (mean inter-item r = .55; 3 characteristics, α = .78), selectivity factor (mean inter-item r = .65; 9 characteristics, α = .94), size and financial need factor (mean inter-item r =.74; 11 characteristics, α = .97), cost factor (mean inter-item r = .52; 3 characteristics, α = .76), homogeneous campus life (mean inter-item r = .51; 4 characteristics, α = .81), and math and science preparation (mean inter-item r = .63; 2 characteristics, α = .77).

Two additional characteristics reflective of diversity in the student body were also used: percentage of female students and percentage of traditionally disadvantaged minority students (i.e., African American, Native American, and Hispanic students). These two characteristics did not hang with the seven institutional factors or each other. Some of these nine characteristics were not reported for all of the schools, but were necessary to calculate factor scores. For these cases the missing data was imputed using an expectancy-maximization algorithm. This algorithm applies an iterative procedure based on all other available pieces of information.

Statistical artifact corrections

Criterion unreliability corrections. Unlike most other studies in this domain, individual grades for each course taken by students during their first year were in the dataset. This allowed for the calculation of the reliability of freshman GPA based upon individual course grades rather than correlating grades from the first and second semester as in other work (e.g., Powers, 2004). Reliability was conceptualized as the mean intraclass correlation (ICC) of individual course grades weighted by course credits. This has the advantage of avoiding issues related to systematic differences across semester (i.e. differences in course difficulty). The ICC can be thought of as representing the correlation between grades in two randomly chosen courses. Since students take different courses, the course coded as "course 1 in 1st semester" or "course 2 in 1st semester" differs for each student at a given school. Therefore, a vector of course grades for each student is coded. The ICC is a good approximation of this scenario since it involves a comparison of variance within-student relative to variance between students. Further, under this conceptualization the ICC does not underestimate reliability since two courses are treated as being randomly chosen (McGraw & Wong, 1996).

Further, the ICC was weighted by course credit in order to mirror how GPA is calculated with higher credit courses having greater weight. For example, the ICC for two four-credit courses was given a weight of 8 (4 + 4) while the correlation between a two-credit course and a three credit course was given a weight of 5 (2 + 3). Zero credit courses were excluded from this calculation because they behaved differently from the rest of the courses. Specifically, (1) zero credit courses typically are not included in calculation of overall GPA; the majority did not receive separate grades (e.g., a course with a required lab section), (2) zero credit courses were rare constituting less than 5% courses taken during the freshmen year, and (3) grades for zero credit courses did not correlate highly with higher credit courses compared to the correlations between grades of high credit courses with each other. In a few instances, negative ICCs result between particular pairs of courses. To correct these estimation errors, Pearson correlations between pairs of courses were used as a substitute.

The Spearman Brown prophecy formula was applied to the course-credit weighted grand mean ICC for each institution to produce a reliability estimate which represented the average number of courses taken by students in their first year for that particular institution. The disattenuation formula was used to correct for criterion unreliability at each institution; the HSGPA-FGPA correlation for each institution was divided by the square root of the corresponding criterion reliability. Corrections for criterion unreliability were employed before corrections for range restriction.

Range restriction corrections. Range restriction refers to reduction in the variance of predictors due to the use of said predictors in the selection of the sample used in a study (e.g., computing HSGPA-grade correlations in samples where HSGPA was used in the selection process). Restricted variance causes a lower HSGPA-FGPA correlation than would occur if in an applicant sample (Sackett & Yang, 2000). To correct this issue, a multivariate range restriction correction was employed, with HSGPA, SAT, and SES as selection variables. Each of these variables could influence whether a student will enroll at a given school, either by the college/university using them in the selection process and/or by the student's decision on the school to select based upon their standing on these variables.

In order to carry out these corrections the unrestricted means, SD's, and correlations among the variables are needed. Two distinct sources were obtained to do so. First, the entire population of individuals taking the SAT and completing a questionnaire reporting SES and HSGPA in 2006 (over 1.25 million students) was used to provide means, SDs, and correlations between the SAT, HSGPA, and SES. This data allowed for an estimate of the entire population. Second, estimates of the means, SD's and correlations in the applicant pool for each specific college or university were desired to have unrestricted estimates of the correlation of interest among each school's unique applicant population. These data were unavailable. However, the schools to which each student sent their SAT scores were available. This provides a reasonable estimate of the school-specific applicant pool and was also a strategy employed by Sackett et al. (2009). Therefore, multivariate range restriction corrections were performed with both the school-specific estimates of the applicant pool and the entire SAT taking population as the referent population (Sackett & Yang, 2000).

Multivariate range restriction corrections are preferred due to their increased accuracy as compared to univariate corrections (Held & Foley, 1994). Means, SDs, and correlations were available for SAT scores, HSGPA, and SES in the restricted and unrestricted populations in each school and at the national level. A multivariate range restriction correction for these three variables was employed, rather than just on the HSGPA, for greater accuracy.

Neither of the populations used for corrections (school level and national) are necessarily preferable. Rather, they each address different yet important questions. Students' SAT scores, HSGPA, and SES likely influence the schools to which they apply. Therefore, SAT, HSGPA, and SES variability may be smaller within any given college's applicant pool than in the total population of college applicants. Using individual school applicant pools for corrections estimates how well SAT and HSGPA could be expected to predict grades within the average college's applicant pool. Corrections to the national level show how well SAT and HSGPA would predict grades if SAT and HSGPA variance was not reduced by students self-selecting into certain colleges' applicant pools. Since both of these questions have merit both types of corrections were applied for comparison purposes.

Data analysis plan

First, the correlation between HSGPA-FGPA was corrected for unreliability and range restriction (both school specific and national) for each institution in the database. To address research question one the extent to which variability exists across college/university will be assessed. If sufficient variability exists, research question two will be examined. Specifically, the nine institutional characteristics categories will be used to predict variability in HSGPA validity across institutions for both the school-specific applicant population and the national population.

Hierarchical linear modeling (HLM) was employed as an alternative approach. HLM allows variance in outcome variables to be analyzed at multiple hierarchical levels and as such is designed to examine nested data (Nezlek & Zyzniewski, 1998). HLM is well-suited to this dataset since it consists of students nested within colleges/universities. In the current dataset HLM can analyze differences between institutions and dependencies in the data such as students within a school being more similar on FGPA scores. In the HLM model employed, using grand-mean centering, HSGPA was a level 1 predictor of FGPA and the nine institutional characteristics were level 2 predictors of the HSGPA-FGPA slope. HLM has the advantage of greater power to detect statistical significance since power is derived from both level 1 and level 2. Further, HLM can handle unbalanced designs (i.e. unequal sample sizes in groups) better because it accounts for differential reliability based on number of level 1 units (i.e. number of individuals within a school).

However, HLM does not allow for statistical corrections. Therefore, this research chiefly focuses on the approach of regressing corrected validity coefficients on institutional characteristics. In this approach the extent to which statistical artifacts explain variance can be examined (e.g., range restriction, sampling error, differences between schools in reliability of grades).

Results

Substantial variability exists in uncorrected HSGPA-FGPA relationship. Validities range from .19 to .58 with a mean of .39 (sd = .09). The mean school-specific range restriction corrected validity across institutions is considerably larger (r = .51; sd =

.08) yet variability still remains (range from .25 to .70). Similarly, the mean national range restriction corrected validity across institutions is .59 (sd = .07) with a range from .36 to .75. After corrections both the national and school-specific validities still contain meaningful variation. Further, it is important to note that the relationship is stronger when correcting the HSGPA-FGPA for the entire test-taking population rather than for the applicant group of a specific institution.

Criterion unreliability within school was calculated from the average reliability of GPA based on course grade ICC. This resulted in a mean ICC across schools of .34 (sd = .05; range = .21 – .49) and number of courses (mean = 10.49, sd = 1.97, range = 4.10 to 21.73). After applying the Spearman Brown Prophecy formula the reliability of FGPA ranged from .71 to 0.92 with a mean of .84 (sd = .04). Like previous work by Munday (1970) which found .80 for the reliability of freshmen, this method suggests that FGPA is fairly reliable within institutions.

There was a great deal of variation in the amount of range restriction on SAT and HSGPA by institution, but generally some restriction of range was found for each school. SAT U-ratios for the school specific applicant pools ranged from .60 to .98 with a mean U-ratio of .82 (sd = .07). SAT U-ratios for the national test-taker population showed a bit more range restriction ranging from .51 to .82 with a mean U-ratio of .66 (sd = .06). HSGPA U-ratios were similar to that of the SAT. For school specific applicant pools the U-ratio ranged from .63 to 1.01 with a mean of .86 (sd = .08). Again, that of the national test-taker applicant pool was more restricted ranging from .46 to 1.02 with a mean U-ratio of .72 (sd = .11). Conversely, there was generally little range restriction on SES when looking within school. SES U-ratios using school specific applicant pools range from .82 to 1.13 with a mean of .97 (sd = .06); SES U-ratios using the national test-taker applicant pool ranged from .24 to .49 with a mean U-ratio of .36 (sd = .04). In the current study, multivariate range restriction corrections were used to simultaneously take into account range restriction on all three of these variables.

Table 2 shows a correlational matrix of the institutional characteristic clusters and HSGPA validity. The results indicate some trends among the characteristics. For example, selective institutions tended to have a higher cost of attendance, provide a more homogeneous college experience, have higher expectations regarding mathematics and science preparation, and have a lower percentage of female students on campus. Since selectivity, cost, and homogeneous campus life were so strongly correlated, the predictive power of a second order composite of these factors was examined as well. In addition, more expensive schools more heavily relied on nontraditional assessment methods in their admissions process. Larger schools with a high percentage of students with financial need were commonly lower in cost of attendance, had students with more varied college experiences, and relied more strongly upon traditional selection tools.

Table 3 lists the multiple regression results for institutional characteristics predicting school-specific corrected validities. In this case, institutional characteristics account for 20% of the variance in test validity across schools. Specifically, results indicated that schools that offered a more homogenous campus experience ($\beta = .36$, $p < .05$) and those with a smaller percentage of disadvantaged minority students ($\beta = -.23$, $p < .05$) tended to have higher HSGPA validities. Table 4 shows the multiple regression results with the second order composite (selectivity, homogenous campus life, and cost) for school-specific corrected validities. This resulted in a somewhat different pattern of results. Percentage of disadvantaged minority students ($\beta = -.26$, $p < .05$) was still significantly related to validity. However, the second order composite ($\beta = -.09$, $p = .37$) was not significantly related to validity suggesting that the unique variance of the homogenous campus life factor drove results. Further, unlike in the model with all nine institutional characteristics, the alternative predictors factor ($\beta = -.24$, $p < .05$) and the size and financial need factor ($\beta = -.26$, $p < .05$) were significant.

National test-taker population range restriction correction validities

Table 5 shows the parallel analyses for the corrected to the national test-taking population. Institutional characteristics predicted slightly more variance ($R^2 = .26$) when correcting to this population. Similar to school-corrected validities, schools with a larger percentage of traditionally disadvantaged students ($\beta = -.27$, $p < .05$) generally showed lower levels of validity. In addition, smaller schools with students with less financial need ($\beta = -.26$, $p < .05$) and schools that relied less heavily on alternative predictors for admissions ($\beta = -.30$, $p < .05$) showed higher validities. Unlike for the school-specific corrections, similar results were obtained when the second order composite was substituted into the regression equation. The size and financial need factor ($\beta = -.25$, $p < .05$), alternative predictors factor ($\beta = -.31$, $p < .05$), and the percentage of disadvantaged minority students ($\beta = -.29$, $p < .01$) were all still significant predictors. Further, the second order composite ($\beta = .22$, $p < .05$) was also a significant predictor.

HLM results

First, in the unconditional model, the intra-class correlation was .117 indicating that roughly 12% of the total variance occurs between schools. HLM results using the nine institutional characteristics are provided in Table 7. These results were slightly different than the corrected results. Unlike previous results less selective schools and lower cost schools tended to have higher validities. However, it is possible that these results are in part driven by range restriction effects since statistical artifacts cannot be corrected for in HLM. The selectivity factor showed a moderate relationship (r = -.35) with HSGPA U-ratio meaning that more selective institutions tended to be more restricted. Table 8 shows results using the second order composite. These findings match the results for the nationally corrected sample. The second order composite, size and financial need factor, alternative predictors factor, and the percentage of disadvantaged minority students were all significant predictors.

Discussion

Study one aimed to examine the extent to which variability in the relationship between high school grades and college performance could be attributed to institutional characteristics. First, substantial variability was found in the validity of HSGPA even after applying corrections for range restriction and criterion unreliability. After correcting for statistical artifacts the standard deviation for HSGPA-FGPA relationship was .08, resulting in a .32 difference between a school 2 SDs below the mean and one 2 SDs above. Absent further information it is difficult to provide an institution with more information than the meta-analytic mean with a relatively wide credibility interval around it.

The current study found that this estimate could be improved upon through the use of institutional characteristics. However, results differed slightly depending on the type of correction used. When looking at the school-specific corrected validities, schools with a more homogenous campus life and those with a smaller percentage of disadvantaged minority students tended to have higher validities. In addition, if a second order composite of selectivity, cost and homogeneity was used smaller schools showed greater validity. Turning to results based upon the corrections to the national test-taking population, smaller schools, schools that relied less on non-traditional selection methods, and those with a lower percentage of minority students generally had higher HSGPA validities. The second order composite related to greater validity as well. Specifically, more selective, higher cost, and more homogenous schools tended to have higher validities. These findings were mirrored when HLM was employed.

Unlike Shen et al.'s (in press) work examining variability in the SAT-FGPA relationship some differences existed depending on the type of range restriction correction. In their work a more consistent set of institutional characteristics was predictive in both cases. Yet, all of the institutional characteristics that were predictive in the current study were also predictive of variability in SAT validity. Only the percentage of minority students was consistently predictive of variability in the HSGPA-FGPA relationship. However, it is important to note that this finding does not contradict previous research which found no differential prediction for minority students (Schmidt, 1988; Kuncel & Hezlett, 2007; Kuncel & Sackett, 2007).

Like Shen et al. (in press), institutional characteristics were more predictive when corrected to the national population. However, they showed a smaller relationship with variability in the HSGPA-FGPA relationship as compared to the SAT-FPGA relationship. The more consistent pattern and higher predictive power for the SAT could be due in part to its standardization. As noted previously, unlike the SAT, two identical HSGPAs might in fact be very different due to a number of factors such as high school difficulty, the types of courses taken in high school, and grading standards. It is also possible that the different pattern of results by correction type was driven by institutions whose applicant pool greatly differs from the national sample. The type of correction employed would result in greater differences in corrected validities among such institutions.

Institutional characteristics could impact the HSGPA-FGPA relationship for a variety of reasons. The current study found that schools which emphasize alternative predictors in their selection systems show lower validities. The use of such alternate predictors might be indicative of an institutional environment where traits and abilities other than those captured by HSGPA are rewarded. For example, schools that stress the importance of extracurricular activities and unique life experiences in essays might more heavily emphasize experiential learning in the structure of their coursework and grades. It was also found that more homogenous schools tended to have higher validities. Higher validity for more homogenous schools could be due to a decreased impact of factors unrelated to HSGPA such as having a part-time job or living off-campus. These factors could diminish time available for study and lower FGPA. Several studies have found that having a part-time job can be detrimental to success in college (Stern & Nakat, 1991; King & Bannon, 2002).

In addition, institutional characteristics could be indicative of systematic differences in the nature of the grades across institutions. Even though grade reliability was corrected for in the current study, construct differences in grades could exist across schools. In other words, grades at different schools may be composed of different reliably measured elements. Because of this, the determinants of performance may differ by school (see Kuncel, Hezlett, & Ones, 2001). For example, some schools may emphasize active participation more heavily in their grades leading to stronger correlations with variables such as extraversion and dutifulness. Other schools might place greater emphasis on written work leading to a stronger relationship with writing skills. In both cases schools may be reliably assessing performance. However, the correlation between HSGPA and FGPA could be affected by the extent to which different variables (e.g. effort, personality, and ability) play a role in the determination of grades.

Ultimately, the current work finds that institutional characteristics can help explain variability in the relationship between high school grades and college grades. Future work should continue exploring why variability exists. This work can help institutions make better decisions about how to apply and use different sources of information in the college application process.

Study 2: Examinations of Non-linearity in the HSGPA-FGPA Relationship

Examination of the extent to which relationships in psychology deviate from linear relationships is often underutilized. This is somewhat surprising given the prevalence of techniques available to predict non-linear relationships and the capability to easily run these models due to far greater computing power than was available in the recent past. This lack of use might in part stem from the widely accepted and well-supported finding that a linear relationship exists between general mental ability and performance (Hawk, 1970; Coward & Sackett, 1990; Arneson, 2007). However, recent evidence in other domains suggests that the examination of non-linear relationships can be fruitful. For example, work in the realm of personality research has shown some success with non-linear models. LaHuis, Martin, and Avis (2005) found support for an inverted U-shape describing the relationship between conscientiousness and supervisory ratings of job performance among government clerical employees. With a sample of college students, Cucina and Vasilopoulos (2005) investigated curvilinear relationships between Big Five personality and academic performance in the first semester. Openness to experience had a significant quadratic but not linear term while conscientiousness had significant linear and quadratic components showing an inverted U relationship. Benson & Campbell (2007) found evidence for non-linear relationships between derailing/dark side personality composites and leadership performance showing a concave downward curve.

Questions of linearity are critical, particularly for grade predictor and criteria. Grades naturally have ceilings that are likely to be met by a non-trivial number of students. This is a direct contrast to scores on many standardized tests (e.g., the SAT). Although there is technically a ceiling, the number of test takers to reach the highest score is so low that the test is without a meaningful ceiling for most settings (for a possible exception see Park, Lubinski, Benbow, 2007). With grades non-linearity could occur because people pile up at the ceiling for the predictor, criterion, or both. This issue is ameliorated to a degree for high school grades which often include the possibility of honor points for honors or advanced placement courses. This raises the ceiling and permits more differentiation among college applicants. Yet, work by Noble and Sawyer (2002) suggests that non-linearity may exist. They used logistic regression and predicted different levels of success in college (GPAs of 3.0, 3.5, 4.0). They found that HSGPA was less useful than an ACT composite when predicting higher grade point averages in the first year of college. In contrast, HSGPA was the better predictor for lower levels of achievement (FGPAs between 2.50 and 3.00). This work suggests that HSGPA might not operate equally across the entire range of FGPA. Further, it may function differently than standardized tests such as the ACT and SAT. In high school, performance may be more related to non-cognitive factors such as effort, attendance, conformity, and motivation. Earlier research by Goldman and colleagues found that the non-achievement components of college grades were associated with average performance (Goldman & Hewitt, 1975; Goldman, Schmidt, Hewitt, & Fisher, 1974; Goldman & Widawski, 1976). High college grades, on the other hand, were more likely to be due to high cognitive ability. These findings suggest that when predicting college grades, HSGPA may level off at higher college GPAs.

Research Question 1: To what extent does the relationship between HSGPA and FGPA deviate from strict linearity? If deviations exist, are they meaningful?

Deviations from linearity in the HSGPA-FGPA relationship could also be driven in part by the nature of FGPA. As noted previously, Ramist, Lewis, and McCamley (1990) found that large differences exist by course type. These differences could in part reflect different levels of cognitive load. Courses that were graded the easiest such as physical education, arts classes, and education likely demand relatively lower cognitive loads. In comparison, the most harshly graded courses were in the hard sciences that are likely to be more cognitively demanding. Therefore, the shape of HSGPA's relationship with college grades may vary due to differences in grading standards and the cognitive load of different course types.

Research Question 2: Does the shape of the relationship between high school GPA and college grades vary by course type?

More recent research suggests that for certain subgroups of the college going population traditional predictors might operate differently. When looking at English as a Second Language (ESL) students, Stricker (2004) found evidence of non-linearity in the relationship between the GRE-Verbal section and TOEFL scores. Specifically, low TOEFL scores showed little relationship with GRE-Verbal scores. However, at the upper range of the distribution TOEFL scores were more strongly related to GRE-Verbal scores. This work suggests that traditional predictors such as the HSGPA and the SAT may operate differently depending on students' language proficiency. In addition, as noted earlier Shen, Kiger et al. (2011) found that differences among language groups can vary by race/ethnicity. Therefore, the current study will conduct analyses within these subgroups as well.

Research Question 3: Does non-linearity in the HSGPA-FGPA relationship vary by language proficiency within racial/ethnicity subgroups?

Expanding on the work of study one, the impact of non-linearity on variability in the HSGPA-FGPA relationship will be examined. The weaker linear relationships between HSGPA and FGPA in some schools could in part be due to non-linear relationships.

Research Question 4: Do schools that show weaker HSGPA-FGPA relationships display greater deviations from strict linearity?

To investigate non-linearity in the HSGPA-FGPA relationship several techniques were employed. First, in order to graphically examine the nature of these relationships a local fitting technique called locally weighted scatterplot smoothing (LOESS) was utilized. The use of smoothing by local fitting has long been employed in the examination of time series plots in which data was smoothed by local fitting of polynomials (Macauley, 1931). This technique was later refined to employ regression analysis (Cleveland, 1979). LOESS effectively combines linear least squares regression and non-linear regression by fitting models to segments of the data and then smoothing the lines to produce a line representing the nature of the relationship whether it be curvilinear or straight (Cleveland & Devlin, 1988). Unlike many other methods with LOESS one does not need to specify a function to fit a model to all of the data in the sample. However, unlike other techniques, such as linear regression, this does make it difficult to express a model that can be easily represented mathematically. Another limitation is its need for large, densely sampled datasets relative to other least squares methods. In the current study this issue is remedied by the use of a large College Board Consortium dataset. LOESS has seen limited use in some areas of psychology such as the modeling of personality disorders (Aluja, Cuevas, García, & García, 2007; O'Connor, 2005), suicide rates (Yamasaki et al., 2004; Partonen et al., 2004) and genetic mapping (Duffy, 2006). However, this technique has seen little application in the areas of educational and industrial/organizational psychology.

Two regression-based procedures were also employed to look for non-linearity: the power polynomial approach and segmented regression. The power polynomial approach is a commonly used technique to ascertain departures from linearity (Coward & Sackett, 1990; Cucina and Vasilopoulos, 2005; Benson & Campbell, 2007). It employs hierarchical regression analyses with a squared predictor term added in the subsequent step; higher terms (i.e. cubed) can follow as well. If a quadratic term shows a significant increment to R^2 over the simple linear model then there is some evidence for non-linearity. Segmented regression (Arneson, 2007) runs regression equations within sections of the data. This allows for more fine-grained comparisons of the slopes in different parts of the data.

Method

Sample

The College Board dataset described in study one was employed in study two as well.

Measures

The following measures were employed in addition to the measures noted in study one:

Demographics. Race/ethnicity was obtained from questionnaires completed by students when taking the SAT. Students indicated what racial/ethnic group they belonged to within eight classification options: (1) American Indian or Alaskan Native; (2) Asian, Asian American or Pacific Islander; (3) Black or African American; (4) Mexican or Mexican American; (5) Puerto Rican; (6) Other Hispanic, Latino, or Latin American; (7) White; and (8) Other Ethnicity. Those who self-identified as Mexican or Mexican American, Puerto Rican, and Other Hispanic, Latino, or Latin American were combined into an overall Hispanic or Hispanic American group in order to preserve an adequate sample size for analyses.

Best Language Spoken. Best language spoken was also obtained from questionnaires completed by students when taking the SAT. Each student indicated their best language spoken from three classification options: (1) English Only, (2) English and Another Language, and (3) Another Language.

Analysis plan

In order to have an adequate sample size to test for deviations from linearity throughout the score range individuals were pooled across institutions. For this reason, unlike in study one, analyses were not run within individual colleges/universities. Further, multilevel analyses were not undertaken (i.e., individuals nested within institutions).

Adjustments to FGPA. Because individuals are pooled across institutions FGPA was adjusted in order to account for differences across school. To accomplish this adjustment, college grades were first standardized within university, then regressed on HSGPA, SAT, and 48 dummy coded variables representing the 49 institutions. The dummy variables' coefficients provided the corrections to be applied to students' GPAs at each university.

Further, to test research question two, separate GPAs were examined within three course categories: humanities, social sciences, and STEM fields.

Subgroup analyses. To address research questions three and four, analyses were conducted within various subsets of the data. In order to examine research question three, additional analyses were conducted within subgroups by race/ethnicity and language proficiency.

Non-Linearity Analyses

LOESS Smoothing. To investigate non-linearity in the HSGPA and FGPA relationship several techniques were employed. First, LOESS was employed to represent the data graphically. Several steps were first taken to prepare the dataset for the use of LOESS techniques. First, HSGPA was rounded to the nearest tenth to collapse values into a smaller range. Next, the average FGPA was calculated at each tenth of HSGPA (e.g. 3.1, 3.2, 3.3, etc.). Then average FGPAs for levels of HSGPA which were calculated by fewer than ten individuals were excluded from analyses in order to help prevent undue influence by outliers. Next a linear regression model and LOESS model were run in order to compare relative fit. The LOESS model was calculated using the LOESS function in the statistical program R. This function enabled the creation of a graphical representation of the relationship between HSGPA and average FGPA.

A number of decisions need to be made for this function. First, the span or the number of cases in each local regression needs to be chosen, the default span of .75 was used in our function. This value was judged to provide an appropriate level of smoothing. Second, the polynomial for the local regression must be decided, a higher polynomial has less bias but greater variance. Again the default value of 2 denoting a quadratic polynomial was chosen. This was employed because cubic and higher polynomials tend not to improve fit much.

Power Polynomial Approach. Next, a power polynomial approach was employed to examine non-linearity (Cohen & Cohen, 1983; Coward & Sackett, 1990). To accomplish this, hierarchical regression analyses were conducted in which HSGPA and HSGPA squared were added in subsequent steps. If a quadratic term shows a significant increment to R^2 over the simple linear model then we have some evidence in favor on non-linearity. This approach was also utilized at the upper range of the ability distribution (those above one SD above the mean) in order to provide a more fine-grained analysis of non-linearity at the upper tail. This was conducted due to the difficulties in detecting ceiling effects noted by Coward (1989).

Segmented Regression. Finally, a segmented regression (Arneson, 2007) approach was employed to further examine the nature of the curvilinearity. In this approach linear regression analyses were conducted at different sections of test score ranges and then the linear regressions were graphed and compared across the regions. This procedure was employed using three different splits of the data (fourths, sixths and eighths) in order to find an appropriate cut. The data was split such that each segment contains roughly the same number of participants. Therefore, the segments do not cover equal ranges of HSGPA.

Results

Table 9 displays the descriptive statistics for HSGPA, FGPA, and discipline specific GPAs. The mean GPA and standard deviation across discipline are very similar. As shown in Table 10 each of the three discipline specific GPAs correlate highly ($r \approx .80$) with overall FGPA. The discipline specific GPAs correlate fairly highly ($r \approx .60$) with each other as well. The level of adjustment to account for institutional grading differences in GPA was nearly identical across GPA type, see Table 10. Further, the rank order in terms of difficulty was largely the same with some slight differences by GPA type. In other words, institutions that graded harshly in STEM fields also graded harshly in the other fields.

The results of power polynomial analyses are shown in Table 11. Results are displayed for the entire dataset and in a sub-sample of participants with test scores above one SD above the mean. The change in R^2 values with the addition of a quadratic term was significant in each case. However, such significance tests are influenced greatly by the large sample sizes. The change in R^2 is relatively small (.002). At the top end of the distribution the same results were found with a significant but small (.002) increase in R^2. Similar results were found across discipline with STEM GPAs showing a slightly greater (.004) yet still small increase in R^2.

Figure 1 displays the LOESS models for the HSGPA-FGPA relationship. In the LOESS model the relationship between HSGPA and FGPA appears to level off at very high HSGPAs. That is, the slope of the line is steeper up to around a high school grade of 4.0 and then it levels off, retaining a generally monotone form but with a less steep slope. For comparison purposes, the results using linear regression is also provided on Figure 1. Figure 1 also shows the segmented regression for the HSGPA-FGPA. This approach similarly indicates that at higher HSGPAs the slope of the regression lines become steeper until leveling off at the top of the distribution.

Due to its ability to clearly illustrate linearity differences across the data range and its similarity with LOESS findings, the results for only segmented regression are shown when broken down by GPA type and ethnicity. Figure 2 shows the results for the three specific disciplines: humanities, STEM, and social sciences. The segmented regressions for humanities and social science GPAs are very similar to that of overall FGPA. Results for STEM GPAs are similar as well but have a slightly flatter line at the lowest HSGPAs.

Table 12 lists the descriptive statistics for HSGPA, FGPA, and the discipline specific GPAs within ethnicity. The means and standard deviations for the White and Asian groups are very similar. In line with previous research, Black and Hispanic students have slightly lower mean values but similar standard deviations when compared to White and Asian students (Roth et al., 2001). Figure 3 shows the results by ethnicity. The segmented regressions of the White and Asian group closely resemble the overall. This pattern holds for these two ethnicities when looking at discipline specific GPAs as well, see Figures 4 and 5. However, for Hispanic and Black groups a more linear pattern is found, see Figures 6 and 7. Looking within discipline for Black students, Figure 6 shows that humanities GPA has a pattern like that of the overall group with a leveling off at the upper ranges of GPA. Conversely, in the STEM and social sciences fields a more linear pattern is found with the slope at the lower range of HSGPA more closely matching the slope at the upper range. Within discipline segmented regression results for Hispanic students were fairly linear in the humanities and STEM fields. Social sciences, on the other hand, displayed steeper slopes in the middle ranges of HSGPA and flatter slopes at the extremes.

Language group was also examined as a potential source for different patterns among Asian and Hispanic students. Table 13 shows the descriptive statistics for HSGPA, FGPA and the discipline specific GPAs broken down by language group. Asian students whose best language was English had very similar but slightly higher mean GPAs than Asian students who were Bilingual or spoke a language other than English best. Similarly, Hispanic students whose best language was English had slightly higher mean GPAs than Bilingual Hispanics. Hispanics who spoke a language other than English best had higher GPAs than other Hispanic students; however, very few students were in this category and it was dropped from subsequent analyses.

Figures 8-10 display segmented regression results for the Asian language groups. The English only language group results were similar to that of the overall and White groups with a slight leveling off at the upper ranges of HSGPA. The bilingual language group also displayed a similar pattern; however, the lower range also was flatter as well.

The most striking differences were found for students whose best language was other than English. For these students the pattern was much more linear throughout the range. Figures 11 and 12 show the results for Hispanic students whose best language was English and those who were best in English and another language respectively. As noted above, results were not examined for Hispanic students who spoke a language other than English best due to an insufficient sample size. The English best group and bilingual group displayed similar results with a fairly linear pattern in each case except with social sciences as the criteria. The segmented regression for the social sciences showed a flatter slope for the extreme values and a steeper slope in the middle.

Expanding on the issues presented in study one, differences in terms of non-linearity were examined among high and low validity schools. The validities for HSGPA and adjusted FGPA ranged from .31 to .63 with a mean of .49 and standard deviation of .08. Schools with a validity 1.5 standard deviations below the mean were selected for the low validity group and those 1.5 standard deviation above the mean were selected for the high validity group. This resulted in a sample of five schools for the high validity group and four schools for the low validity group. Figure 13 displays the results for both the high and low validity groups. The high validity group displayed a more strictly linear pattern with some flattening for students with lower HSGPAs particularly for STEM coursework. The low validity group displayed a slightly different pattern typically showing a flatter relationship at the lower end of HSGPA and a slightly flatter slope at the upper end.

Discussion

The use of techniques beyond linear models can greatly aid our understanding of complex relationships. The current study aimed to show that there are benefits in looking beyond linear regression in the examination of the relationship between high school GPA and freshman GPA. The power polynomial approach found little evidence of departures from linearity, however, the other two techniques applied showed some evidence in support of departures from strict linearity. The LOESS model indicated a leveling off at the top end of the high school GPA distribution. In addition, using segmented regression there was also evidence of a leveling off at the upper end of the distribution. Therefore, concerning research question one meaningful deviations do appear in the overall relationship between HSGPA and FGPA.

Research question two posited that differences might exist due to academic discipline. However, little evidence was found in support of major differences among the course types examined. Generally speaking, a similar pattern was found regardless of course type. However, it is important to note that relatively broad groupings of coursework were employed. Future research which can employ more finely discriminated course groupings would be beneficial.

Turning to research question three, the extent to which ethnicity and language group played a role was also of interest. Some differences did exist based upon these categories. Broadly speaking, the pattern for White and Asian students generally mirrored that of the overall group with increased slope in the middle ranges of HSGPA and some leveling off at the upper regions. The pattern for Black and Hispanic students, on the other hand, tended to be more strictly linear. Since, these groups are overrepresented in lower SES groups these findings could in part be influenced by school quality and social class. Schools with a lower achieving population and schools with an overall lower SES might less frequently provide AP and college preparatory coursework which allow for grades above 4.0. Therefore, due to limited availability, the Black and Hispanic students that did take such coursework might be significantly better students than their peers. In schools were such coursework is more ubiquitous this might not be the case. Study three attempts to address the impact of some of these SES issues.

Language group further complicates the picture. For Asian students the English Best group and the Bilingual group produced results that mimicked the Asian group as a whole. However, Asian students whose best language was other than English tended to show a more linear pattern across the entire range of HSGPAs. For these Asian students it is likely that only the strongest students take advanced coursework that offer grades above 4.0 causing a more linear pattern. The pattern for Hispanic students whose best language was English, similar to the overall Hispanic group, displayed a fairly linear relationship. Results for Bilingual Hispanic students were less clean but typically showed very steep slopes for HSGPAs between 3 and 4 and relatively flatter slopes at the extremes. Insufficient data was available to assess Hispanic students whose best language was other than English.

Relating to the issues of variability in validity examined in study one, research question four posited that non-linearity might be the source of some of these validity differences. There is some evidence for non-linearity playing a role in validity differences; low validity schools display an overall flatter relationship with less steep relationships at both the upper and lower segments for HSGPA. This effect is particularly pronounced for FGPA and social sciences GPA.

Overall, White and Asian students tended to follow the pattern of an increasingly steep slope in the middle range of HSGPAs followed by a flattened slope at very high levels of HSGPA. Notably, Hispanic students, Black students, and students whose best language was not English displayed a pattern more closely following strict linearity. These differences could be a result of different high school course-taking patterns and/or the differential availability of high school coursework with weighted grades. However, due to smaller sample sizes results within ethnicity should be interpreted less strongly. Future research should examine these issues further to more fully test this phenomenon.

In addition, in psychology as a whole we need to consider applying non-linear techniques more often in order to better understand the nature of important relationships. As noted by Coward and Sackett (1990) departures from linearity can have a significant effect on the impact of certain types of selection decisions. If a linear relationship exists then the use of a minimum cut-off could cause a large drop in performance. If, however, the relationship is non-linear and levels off at higher scores, if a cut-off is used where the relationship is less steep this drop in performance would be diminished. The current work suggests that such a relationship might exist with HSGPA.

These results also raise questions about the joint relationships of multiple predictors within a selection of admissions system. If predictors demonstrate plateaus (as does HSGPA) then greater weight should be given to other predictors within those score ranges. This adds a potentially interesting layer of complexity and sophistication to selection systems.

Study 3: The Impact of SES on the HSGPA-FGPA Relationship

SES is one of the most studied variables in educational research (Sirin, 2005). Typically, it refers to an individual's or a family's ranking according to access to valued commodities such as wealth, power, and social status (Mueller & Parcel, 1981). However, educational research commonly operationalizes SES as a composite of mother's education, father's education, and parental income (Sackett et al., 2011; Sirin, 2005). Despite widespread concern, SES has typically shown a fairly weak relationship with college grades; Sackett et al. (2011) found a relationship of .13 with FGPA. Further, as noted above, including SES in a model predicting FGPA does not substantially reduce the importance of the SAT and HSGPA. The lack of importance for this relationship is somewhat surprising given larger values found when looking at academic performance earlier in life. When looking at grades K-12, Sirin's (2005) meta-analysis found a moderate relationship (r = .28) between SES and academic performance.

However, the impact of SES on the relationship between HSGPA and college performance might be more important when looking beyond simple correlations. SES might cause non-linearity in the HSGPA-FGPA relationship. For example, the difference between having parents that complete college compared to parents that complete a graduate degree might be less important relative to the difference between completion of a high school degree and a college degree. Further, at lower income levels a difference of $10,000 dollars might be substantial while at higher levels it might be relatively unimportant. A few recent studies have shown these types of curvilinear effects for neighborhood SES on academic achievement for children (Carpiano, Lloyd, & Hertzman, 2009; Dupere, Leventhal, Crosnoe, & Dion, 2010). In addition, the typical practice of combining SES variables in a unit-weighted composite might obscure these effects.

Research Question 1: Is the relationship between HSGPA-FGPA affected by SES? Do different components of SES have similar effects?

In addition, research suggests that SES can be an important indicator of future educational choices and performance. A great deal of research points to its critical role in early childhood development, most prominently in language development. Numerous studies have found that higher SES mothers talk more to their children than do lower SES mothers (Hoff, Laursen, & Tardif, 2002; Hart & Risley, 1995). The magnitude of these differences is staggering; when looking at families with a child under 2 years old, Hart and Risley found that over the course of one week children of high SES parents hear 215,000 words, children of middle SES parents hear 125,000 words, and children of lower SES parents who are on public assistance hear 62,000 words. Further, these differences appear to have an impact on vocabulary. At age 3, the vocabulary of higher SES children was roughly double that of lower SES children. Evidence in support of critical periods for language development early in life suggests that such differences could have a lasting impact (Ruben, 1997).

SES can also play an important role in terms of educational aspirations (Hearn, 1984; Karen, 1991; Kao & Tienda, 1998; Lareau, 1987; Walpole, 2003), and afforded opportunities in the curricula (Garibaldi, 1997; Matthews, 1984; Orfield, Eaton, & Jones, 1996). Horn and West (1992) found that parental expectations of their children's education related to students' academic success, behavior, self-concept and motivation to learn. Similarly, data from the National Center for Education Statistics' National Education Longitudinal Study of 1988 shows that at high poverty schools 49% of parents of eighth-graders expected their children to finish college, compared to more than 66% of parents of students attending low-poverty schools (Garibaldi, 1997). Given these differences it is possible that low SES students are categorically different from their higher SES peers. The HSGPA of a low SES student might not be equal to that of a higher SES student due to difference in course-taking patterns and/or attendance at a less rigorous high school. Lower aspirations and expectations could also lead to differences in the meaning of college grades. If, for example, low SES students disproportionately select easier majors and courses their FGPA could be inflated relative to students who took more difficult coursework.

Research Question 2: Are low SES students systematically different from their higher SES counterparts in terms of academic readiness and course-taking patterns?

In addition, it is important to note that minority students are disproportionately represented in low SES groups. High-poverty urban, suburban, and rural public schools are more likely to have larger minority student populations than low-poverty schools (Garibaldi, 1997). Therefore, race/ethnicity also merits attention in the study of the impact of SES. It is possible that the effect of SES on academic performance differs by subgroup. For example, the importance of education and learning might vary by subgroup (Matthews, 1984). Due to likely differences in academic opportunities, preparedness, and aspirations for low SES students, the current study aims to examine the impact of SES on the HSGPA and FGPA relationship.

Research Question 3: Does the impact of SES on the HSGPA-FGPA vary by race/ethnicity?

Finally, as noted previously, gender might play a role as well (Burton & Ramist, 2001). Burton and Ramist found that important differences exist in the course-taking patterns of men and women. Further, recent trends show that more women are attending college and earning degrees than men. In 2008–09 women earned 57.2% of Bachelor's degrees (NCES, 2011). The current study seeks to determine if SES plays a role in this trend and if relationship between HSGPA and FGPA is impacted by gender.

Research Question 4: Does the impact of SES on the HSGPA-FGPA vary by gender?

Method

Sample: The College Board dataset described in study one was employed in study two as well.

Measures

In addition to the measures already noted in study one the following measures were employed:

Student Descriptive Questionnaire (SDQ). The SDQ is a 331-item self-report measure completed when students take the SAT. Variables of interest gathered from the

SDQ in the current study include academic rigor (e.g., number of courses taken, honors and advanced placement courses taken), academic aspirations (e.g., degree goal and intended major) and demographic information (e.g., father's and mother's education and family income, ethnic and gender identification, and ESL status).

SES. Three SES variables were provided from the SDQ: father's education, mother's education, and family income (the natural log of family income was used). In addition, a unit weighted composite of these three variables was created to represent SES. In the entire population of SAT takers the correlation between the father's education and mother's education was .60; father's education was slightly more related ($r = .42$) to family income than mother's education ($r = .39$).

Analysis plan

First, to address research question two, the nature of the low SES college student will be examined to determine if they are in fact distinct. Intended majors, course-taking patterns, and rates of intentions to ask for help in college will be examined by SES. Since, the current dataset includes only first year course data; rates in specific rigorous courses will be examined: chemistry and calculus. Students taking these courses in their first year of college are expected to possess higher levels of academic preparedness and aspirations as compared to the typical student. Further partial correlations between SES and indicators that show large differences by SES group membership will be run controlling for ability with the SAT as a covariate.

Next, a regression framework will be employed to examined research question one. Previous research (Sirin, 2005; Sackett et al., 2009) has typically used SES composite as a predictor in regression models; however, this method might obscure results. First, regression analyses with each individual SES component separately included in the model will be performed. Next, to address issues surrounding non-linearity, analyses will be conducted on subgroups based on SES. If different HSGPA-FGPA relationships are shown for different SES groups then there will be evidence that HSGPA does not operate the same across all levels of SES. To test for non-linearity in the HSGPA-FGPA relationship the segmented regression approach described in study two will also be utilized within SES subgroups. As in study two segments are created such that each one includes roughly the same number of participants. Therefore, the segments do not cover equal ranges of HSGPA.

Adjustments to FGPA. As a further examination of research question two, analyses were conducted by course type. Separate GPAs were examined within three course categories: humanities, social sciences, and STEM fields.

Subgroup analyses. To address research question three, following overall analyses, separate analyses were also be performed by race/ethnicity.

Results

Table 14 shows the descriptive statistics for the SES variables, HSGPA, and college grade variables. Some differences were found by ethnicity. White students had the highest mean scores across the SES variables and had the least variability. Asian students had the next highest mean SES scores followed by Black students while Hispanic students had the lowest SES means. As seen in Table 15, values for the national population of SAT-takers generally follow the same overall pattern. However, the national population shows slightly lower mean SES values and slightly greater differences by ethnicity. Bar charts showing the shape of the distribution of the SES composite and the natural log of income are displayed in Figure 14. Both variables are negatively skewed with a piling of data at the higher end of their respective distributions. Figure 15 shows the distributions for father's and mother's education. Again there is some negative skew in the data with most respondent's parents possessing a college degree or higher. A correlational matrix between SES, HSGPA, and college grade variables is shown in Table 16. The SES composite was more strongly related (r = .19) to FGPA than each of its individual components. Father's education was nearly as predictive (r = .18) followed by mother's education (r = .15) and income (r = .11). The SES composite and the individual components related more strongly to FGPA than discipline specific GPAs. Table 17 displays the correlation between HSGPA and FGPA within SES group. The relationship between HSGPA-FGPA was greater for the highest SES students (r = .55) than their lowest SES counterparts (r = .47).

In order to examine differences by SES, students were split into five equal groups by SES composite. Looking at intended majors there were some slight differences, see Table 18. Interestingly, the bottom 20% in terms of SES was the least likely group to intend to major in the humanities (10%) and the most likely to intend to major in hard sciences (44%). The opposite pattern held true for the top 20% in SES; they were the most likely to intend to major in the humanities (16%) and the least likely to group to intend to major in hard sciences (36%). This finding is at odds with the hypothesis that low SES students might be less prepared for college and therefore would be likely to gravitate towards less harshly graded coursework and disciplines. However, these findings could in part be a result of lower SES students being more likely to choose more lucrative majors (Davies & Guppy, 1997; Arcidiacono, 2004). Findings for degree goals were more in line with expectations. The highest SES group was the most likely (66%) to aspire to an advanced degree compared to 58% for the lowest SES group.

As shown in Table 19, there were a few notable differences in course-taking patterns by SES group. Students in the highest SES group took slightly more natural science courses (.18 years more) and social science courses (.18 years more) than the lowest SES group. At the collegiate level, the highest SES group took more humanities credits (.29 years more) than the lowest SES group. In addition, differences arose in terms of percentages of students who took honors coursework in high school and took calculus in college, see Table 20. Across all disciplines at the high school level roughly ten percent more of the highest SES students took honors level coursework as compared to percentages for the lowest SES group. The highest SES group was also more likely to take a calculus course in college (33%) than the lowest SES group (26%). However, as shown in Table 21, after controlling for ability (SAT score), the partial correlation between SES and course-taking patterns (in both high school and college) showed little relationship ($r < .05$) indicating essentially no relationship between course-taking patterns and SES when ability is held constant.

Larger differences existed in terms of intentions to seek help at college and intentions to have a part-time job as shown in Table 19. The SDQ asked students to report intentions to seek help improving academic skills, study skills, and developing educational and career plans. Greater percentages (~10% more) of the lowest SES students reported planning on seeking help in all help seeking categories when compared to the highest SES students. After controlling for ability, the partial correlation between SES and a unit weighted composite of help seeking was slightly negative (r = -.06) indicating a weak tendency for lower SES students to seek more help, ability held constant (see Table 21). Intentions to have a part-time job during college were also more common among the lowest SES students (70%) than for the highest SES students (49%). Partial correlations between SES and intentions to have a part-time job were moderate (r = -.16) with lower SES students being more likely to intend to have a job when holding SAT score constant.

Next, segmented regression for the HSGPA-FGPA relationship was employed within SES group in order to assess differences by social class. As shown in Figure 16, for the overall sample some differences were found by SES group. Regardless of SES, the slope of the relationship tended to increase up to around a HSGPA 4.0 at which point the slope of the line flattened. However, at the lower end of HSGPA, lower SES students showed a flatter slope while higher SES students maintained a more strictly linear pattern up until HSGPAs greater than 4.0. Further this effect was not an artifact of greater representation by lower SES students at the lower end of HSGPA. When regression equations were run with equal endpoints this pattern persisted. Figures 17 and 18 show results split by gender. Gender had little impact on this relationship; the pattern was nearly identical for males and females.

Looking at this relationship by ethnicity in Figures 19 – 22 some differences emerged. However, it is important to note that due to smaller sample sizes results within ethnicity are more exploratory in nature. First, the pattern for White students mirrored that of the overall sample almost identically. Asian students, on the other hand, tended to show a more consistently linear pattern among higher SES students. Lower SES Asian students showed a pattern more similar to the overall group pattern. Similarly, Black students in higher SES groups showed a more linear pattern while those in the lowest group showed a much flatter slope for HSGPAs above 4.0. Finally, the pattern for Hispanics students generally followed that of both the White and the overall group.

Figure 23 shows the segmented regression for the HSGPA-FGPA relationship split by income group. The lowest income bracket shows a fairly strongly linear pattern with some increasing slope in the middle segments. In contrast, the rest of the income brackets display the pattern commonly found for the SES composite with increasing slopes in the middle segments and a leveling off at the highest segment. Again, some differences were found by ethnicity as shown in Figures 24 – 27. The pattern for White students mimicked that of the overall sample. Like for the SES composite, Asian students showed a more strictly linear pattern in higher income patterns and flattened slopes for extreme values in lower SES groups. The patterns for Black and Hispanics students weren't as clean. Generally speaking, Black students showed a fairly linear relationship regardless of income bracket. Hispanic students in higher income brackets tended to display flatter relationships at the highest HSGPAs.

Next, the impact of parent's education was examined on the HSGPA-FGPA relationship. Figures 28 – 32 show the segmented regressions for the overall group and by ethnicity split by father's education and Figures 33 – 37 show these relationships split by mother's education. The pattern for both parental education variables was very similar. Looking at the overall group and that of White students, similar to findings for other SES variables, there was a pattern of slightly increasing slopes until HSGPAs of above 4.0 at which point a flatter slope was found. Asian students generally followed this pattern as well; however, there was more evidence of a flatter relationship at the lower end of HSGPA. Similar to their pattern for other SES variables Black students had a fairly linear relationship except at the lower education level (Less than a college degree) which showed a flattened and even slightly negative slope for HSGPAs above 4.0. In contrast, the pattern for Hispanic students mirrored the overall and White group.

Discussion

Study three aimed to further examine the findings of study two through an examination of the impact of SES on non-linearity in the HSGPA-FGPA relationship. First, the nature of the low SES student was examined. A few noteworthy differences emerged between low SES students and high SES students. Low SES students were less likely to take honors coursework in high school and less likely to take a calculus course in college. However, when controlling for ability there was little relationship between SES and course-taking patterns. Pertaining to research question two, overall course-taking patterns showed little evidence to indicate that low SES students are dramatically different than their higher SES counterparts. Differences that appeared to exist were largely a function of higher ability levels in higher SES groups.

Larger differences emerged in terms of intentions to seek help in academic areas and intentions to have a part-time job. Higher percentages of low SES students planned on seeking extra help to improve academic skills in college across all domains. Further, even after controlling for ability lower SES was related to intentions to seek help. This suggests that low SES students may feel less prepared for college level coursework and/or are more willing to seek help to address this issue. It was also more common for low SES students to intend to have a part-time job than their higher SES peers. Again, this relationship persisted when controlling for ability. While it is increasingly common for students to work part-time, research indicates that it can be detrimental to success in college (Stern & Nakat, 1991; King & Bannon, 2002). Working a part-time job can have many negative effects on the college experience. King and Bannon found detrimental effects on grades, limitations on course choice, and lower course enrollment. In addition, they found that low-income students were more likely to work longer hours further exacerbating these effects.

Turning to the central question, for the total sample, SES had some effects on the relationship between HSGPA and FGPA. Interestingly, SES had little impact on the HSGPA-FPGA relationship for all but the lowest SES students. Generally, lower SES students showed flatter relationships at the low end of HSGPA compared to higher SES students who displayed a more consistently linear pattern up to HSGPAs of 4.0. These findings suggest that SES may be more important for low ability students and that above a certain threshold SES is of little consequence. At low ability levels a higher SES might insulate students from pitfalls. Further higher ability students might be able to overcome such issues regardless of SES. For example, low ability but high SES students might have a better understanding of the expectations at college due to having more highly educated parents. With sufficient ability these expectations could be met regardless of previous familiarity. Low ability and low SES students, on the other hand could be particularly at risk. The overall findings are generally in line with previous work which has found that SES has little impact on the relationship between HSGPA-FGPA (Sackett et al., 2011). Further, the three different indicators of SES (Father's education, Mother's education, and Income) functioned similarly supporting the use of a composite.

Like study two there were some differences in linearity due to ethnicity. However, these results are more exploratory in nature due to smaller sample sizes. White students showed a pattern that was very consistent with the overall sample. This finding held true regardless of SES indicator. When looking within SES group, segmented regression results for Asian students tended to be more strictly linear for higher SES groups with a much slighter reduction in slope for HSGPAs above 4.0. These differences could be a result of many factors. For example, Asian students have been found to have higher educational aspirations and are more likely than other groups to view education as the best means to overcome discrimination and achieve a high social status (Goyette & Xie, 1999). This could result in a more linear pattern due to greater similarities in motivation and performance differences based more predominately on ability. In addition, heterogeneity within the categorization of Asian exists in terms of culture and the nature of their immigration to the U.S. Further, the different subgroups of Asian Americans are not evenly represented in terms of SES. The average family incomes of Japanese, Chinese, South Asian, and Filipino Americans are much higher than Vietnamese, Laotian, Cambodian, and Hmong Americans (Goyett & Xie, 1999). Finally, as shown in study two some differences can arise due to language group. However, due to sample size limitations these effects could not be examined. Differences across SES group could in part be due to all or some of these factors.

Results for Black students also differed depending on SES group. Similar to Asian students, the segmented regression results for Black students was generally more linear for higher SES groups than low SES groups. This could be a result of fewer advanced high school coursework available for low SES students such that higher grades for these students does not reflect similarly higher gains in knowledge or represent higher ability levels.

Unlike the other minority groups, segmented regression results for the Hispanic group tended to closely resemble that of the overall group. This finding is somewhat surprising given there is a great deal of diversity within Hispanic Americans (Arias, 1986) similar to that of Asian Americans. Further, Arias notes that the educational attainment of Hispanic students lags far behind that of most groups in the U.S. with great underrepresentation in higher education.

Ultimately, the current study showed that SES does have some effect on the HSGPA-FGPA relationship. Results indicated that SES may be most important for low ability students and further that above a certain HSGPA threshold SES is of little consequence. In addition, the current study suggests that some differences may occur by ethnicity. For White and Hispanic students little difference was found across SES group. With Asian and Black students some slight differences did occur. The cause of such differences is difficult to determine. Further, the small sample sizes for the minority groups examined make it difficult to draw strong conclusions.

There are many avenues for future research on the role of SES. First, research utilizing more detailed examinations of course-taking patterns could be beneficial. The current study found few differences in course-taking patterns when controlling for ability. However, insufficiently detailed course information was available for examinations beyond a general level in the current work. More fine-grained analyses of the types of coursework taken by students could result in more meaningful differences by SES. In addition, work examining the impact of SES at multiple levels would also be instructive. Both neighborhood and school-level SES could play a role in shaping the academic preparedness of students and impact the predictive power of their HSGPAs. Recent work has shown that aggregate measures of SES can be more predictive than individual level SES (e.g. Carpiano, Lloyd, & Hertzman, 2009; Dupere, Leventhal, Crosnoe, & Dion, 2010). Finally, information on the high schools attended by students was unavailable in this work. Work by the NCES (1984) has shown that significant differences can exist in grading standards across high school. This information would help promote better understanding of variables such as high school difficulty and academic preparedness.

General Discussion

Previous research has established high school grades as a strong and unbiased predictor of academic performance (Burton & Ramist, 2001; Sackett et al., 2011). The use of HSGPA as a predictor is widespread in both educational and occupational contexts. However, in some respects, important aspects of using HSGPA as a predictor of academic performance have yet to be satisfactorily examined. The current work sought to address some of these issues through a more detailed examination of HSGPA across three studies.

One such area relates to the sizable variability in the validity of HSGPA across institutions. In study one these differences across institutions were found to persist even after correcting for statistical artifacts of range restriction and criterion unreliability. Further, it was shown that characteristics of colleges/universities could shed some light on this variability. Improved knowledge for an individual school relating to the usefulness of HSGPA as a predictor has important benefits. Armed with this information, schools can refine and improve their selection criteria. Meta-analytic means can help inform us on the general nature of relationships. However, further work is often needed to more clearly explain relationships. Future work should seek to understand the mechanisms through which institutional characteristics impact validity. For example, examinations on the nature of grades and grading practices across institution merits further attention.

The examination of non-linear relationships is another area of research that is often underutilized. Research techniques in psychology all too often only involve solely linear approaches. Study two sought to apply non-linear approaches to better understand how HSGPA operates as a predictor of college performance across its entire range. This approach proved to be important as deviations from linearity were detected. Overall, the relationship between HSGPA and FGPA showed a pattern of increasing slopes up to HSGPAs of 4.0 at which point the relationship showed a considerable flattening. This general pattern held across ethnicity and course type. However, when looking within language group a more linear pattern was observed for students who spoke a language other than English best. The results of this study indicate that non-linear techniques should be more commonly employed. Further they indicate that caution should be employed when trying to distinguish between students in the extreme ranges of HSGPA. Other variables such as the SAT are likely to be better tools for differentiating between students at HSGPA above 4.0.

Finally, the major issue of social class was addressed in study three. Much attention has been given to this variable. The current study attempted to add to this literature in two ways. First, through an examination of distinguishing characteristics of low SES students and second by employing non-linear techniques to search for effects of SES on the HSGPA-FGPA relationship. The high school course-taking patterns of low SES students differed from their higher SES peers in several important ways. However, these differences proved to be largely a result of differences in ability. Two key differences emerged after taking ability into account. Low SES students were more likely to plan on having a part-time job and they more commonly planned on seeking additional help to improve academic skills upon entry to college. These differences could have important effects on the success of low SES students in college and direct efforts to improve the success of low SES students. Efforts to ensure that low SES students have access to resources for remedial help should be undertaken and institutions should continue to work to provide financial aid in terms of scholarships, loans, and work study. More flexible course-taking opportunities should also be provided.

SES had little impact on the HSGPA-FPGA relationship for all but the lowest SES students. Generally, results mirrored those found in study two. However, low SES students showed flatter relationships at the low end of HSGPA compared to higher SES students who displayed a more consistently linear pattern up to HSGPAs of 4.0. These results suggest that students with a low HSGPA and low SES are particularly at risk. Academic support systems aimed at helping these students successfully transition into college could be beneficial.

Little difference existed through the use of different SES indicators. Some slight differences emerged when ethnicity was considered. White and Hispanic students showed results that were consistent with overall patterns. In contrast, some differences were found for Asian and Black students. Lower SES Asian and Black students displayed a pattern like that of White and Hispanic students. Higher SES Black and Asian students showed a more strictly linear pattern. As noted earlier, these differences could be caused by a multitude of factors such as heterogeneity within ethnicity groups, systematic differences in high schools attended, and/or language and generational differences. Further, results within ethnicity are more exploratory in nature due to smaller sample sizes. Ultimately, further work is needed to more closely examine and replicate such findings.

Although the current work has expanded the literature through an examination of less frequently studied issues much work remains. Specifically, future work could expand on the current findings in three key areas. First, the current study was limited by its inability to adjust for differences by high school. Although information on the coursework taken in high school was examined, data on the high schools attended by students were unavailable. Differences between the same coursework across high schools can be substantial and the impact of high school quality on academic success in college merits further attention. The validity of HSGPA can greatly depend on high school quality (Bassiri & Schulz, 2003).

In addition, more detailed information on subgroups within minority groups could promote greater understanding of differences observed by ethnicity. More information on variables such as language proficiency, generational status, and acculturation could prove useful. Research has found that significant differences exist within the coarse racial/ethnicity categories typically employed (Arias, 1986; Goyett & Xie, 1999). Further, as shown in Shen, Kiger et al. (2011) these variables can be the true drivers of differences between ethnicities.

Finally, information on group level SES variables could support better understanding of the impact of SES. Current research indicates that group level SES such as neighborhood and school-level SES can have important effects on student success (e.g. Carpiano, Lloyd, & Hertzman, 2009; Dupere, Leventhal, Crosnoe, & Dion, 2010). Higher

SES schools differ from lower SES schools in many potentially meaningful ways. For example, higher SES schools have more material and financial resources (Chiu & Khoo, 2005; Tate, 1997), more qualified teachers (Berliner, 2001; Darling-Hammond, 2007), and higher teacher expectations (Rumberger & Palardy, 2005). Some research has even found that group-level SES plays a larger role than individual SES (Sirin, 2005).

In conclusion, the current study supports the value of looking beyond traditional research methods when examining key relationships that may predict academic performance in college. Studying these issues beyond providing a meta-analytic mean and considering patterns other than strictly linearity facilitates greater understanding of these issues. This work may be useful in better predicting academic performance in college and guiding efforts to improve the success of college students who are at higher risk. Future research should more frequently employ these methods to promote better understanding of these issues.

References

Aluja, A., Cuevas, L., García, L. F., & García, O. (2007). Zuckerman's personality model predicts MCMI-III personality disorders. *Personality and Individual Differences, 42*, 1311-1321.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics, 121*, 343-375.

Arias, M. B. (1986). The context of education for Hispanic students: An overview. *American Journal of Education, 95*, 26-57.

Arneson, J. J. (2007). An examination of the linearity of ability-performance relationships among high scoring applicants. (unpublished doctoral dissertation). University of Minnesota –Twin Cities.

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher, 38,* 665-676.

Bartlett, C. J., Bobko, P., Mosier, S. B., & Hannan, R. (1978) Testing for fairness with a moderated multiples regression strategy. *Personnel Psychology*, *31*, 233–242.

Bassiri, D. & Schulz, E. M. (2003). Constructing a universal scale of high school course difficulty. *Journal of Educational Measurement, 40, 147-161.*

Benson, M. J., & Campbell, J. P. (2007). To be, or not to be, linear: An expanded representation of personality and its relationship to leadership performance. *International Journal of Selection and Assessment, 15*, 232-249.

Berliner, D. (2001). *Our schools vs. theirs: Averages that hide the true extremes* (No. CERAI-01-02). Milwaukee, WI: Center for Education Research, Analysis, and Innovation.

Bowen, W. G., and Bok, D. (1998). *The shape of the river.* Princeton, NJ: Princeton University Press.

Breland, H., Maxey, J., Gernand, R., Cumming, T., & Trapani, C. (2002). *Trends in college admissions 2000: A report of undergraduate admissions policies, practices, and procedures*. Research report submitted to ACT, Inc.; Association for Institutional Research; Educational Testing Service; the National Association for College Admission Counseling; and the College Board.

Bridgeman, B. McCamley-Jenkins L., & Ervin, N. (2000). Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning test. New York: College Entrance Examination Board.

Brookhart, S.M. (1998). Why ‘grade inflation’ is not a problem with a ‘just say no’ solution. *National Forum, Phi Beta Kappa Journal*.

Bureau of Labor Statistics. (2011). Occupational Outlook Handbook, 2010-2011 Edition. Retrieved from http://www.bls.gov/oco.

Burton, N., & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980* (College Board Research Rep. No.2001–2). New York: College Board.

Burton, N. W., Morgan, R., Lewis, C., and Robertson, N. (1989). *The predictive validity of SAT and TSWE item types for ethnic and gender groups.* Paper presented at the

1989 annual meeting of the American Educational Research Association and the National Council on Measurement in Education, San Francisco, CA.

Caldas, S. J., & Bankston, C., III. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90,* 269–277.

Callan, P. M. (2008). The 2008 national report card: Modest improvements, persistent disparities, eroding global competitiveness. Retrieved from http://measuringup2008.highereducation.org/commentary/callan.php.

Camara, W., Kimmel, E., Scheuneman, J., & Sawtell, E. A. (2003). *Whose grades are inflated?*. (College Board Research Rep. No.2003-4). New York College Board.

Carpiano, R. M., Lloyd, J. E. V., & Hertzman, C. (2009). Concentrated affluence, concentrated disadvantage, and children's readiness for school: A population-based, multi-level investigation. *Social Science & Medicine, 69*, 420 – 432.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action.* Amsterdam: North-Holland. (Reprinted and revised from *Abilities: Their structure, growth and action,* by R, B. Cattell, 1971, Boston: Houghton Mifflin).

Chiu, M. M., & Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: Country, school, and student level analyses. *American Educational Research Journal, 42,* 575–604.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement, 5,* 115-124.

Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*, 829–836.

Cleveland, W.S. & Devlin, S.J. (1988). Locally-weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association 83*, 596–610.

College Board. (2008). *College Handbook 2009*. New York, NY: Author.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum.

Coward, W.M. (1989). *A statistical power study of tests for curvilinearity in bivariate relationships*. Unpublished master's thesis, University of Illinois, Chicago, Illinois.

Coward, W.M. and Sackett, P.R. (1990). Linearity of ability–performance relationships: A reconfirmation. *Journal of Applied Psychology, 75*, 297–300.

Cucina, J.M. and Vasilopoulos, N.L. (2005) Nonlinear personality–performance relationships and the spurious moderating effects of traitedness. *Journal of Personality, 73*, 227–259.

Davies, S. and Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social Forces, 75*, 1417–1438.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of "No Child Left Behind." *Race Ethnicity and Education, 10*, 245–260.

Dupere, V., Leventhal, T., Crosnoe, R., & Dion, E. (2010). Understanding the positive role of neighborhood socioeconomic advantage in achievement: The contribution of the home, child care and schoold environments. *Developmental Psychology, 46*, 1227-1244.

Duffy, D. L. (2006). An integrated genetic map for linkage analysis. *Behavior Genetics*, *36*, 4-6.

Elliott, R., & Strenta, A. C. (1988). Effects of improving reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement, 25,* 333-347.

Gafni, N., & Bronner, S. (1998). *An examination of criterion-related bias in the testing of Hebrew- and Russian-speaking examinees in Israel*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Garibaldi, A. M. (1997). Four decades of progress…and decline: An assessment of African American educational attainment. *Journal of Negro Education, 66*, 105–20.

Geiser, S., & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *Research & Occasional Paper Series: CSHE.6.07*. Retrieved from http://cshe.berkeley.edu/.

Goldman, R. D., & Hewitt, B. N. (1975). Adaptation-level as an explanation for differential standards in college grading. *Journal of Educational Measurement, 12*, 149-161.

Goldman, R. D., Schmidt, D. E., Hewitt, B. N., & Fisher, R. (1974). Grading practices in different major fields. *American Education Research Journal, 11*, 343-357.

Goldman, R. D., and Widawski, M. H. (1976). A within subjects technique for comparing grading standards: Importance in the validity of the evaluation of college achievement. *Educational and Psychological Measurement, 36*, 381-90.

Goyette, K., & Xie, Y. (1999). Educational expectations of Asian American youths: Determinants and ethnic differences. *Sociology of Education, 72*, 22-36.

Guion, R. M. (1998). Analysis of bias. In Guion, R. M., *Assessment, measurement, and prediction for personnel decisions* (pp. 446-465). Majwah, N.J.: Lawrence Erlbaum Associates.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD, US: Paul H Brookes Publishing.

Hawk, J. (1970) Linearity of criterion-GATB aptitude relationships. *Measurement and Evaluation in Guidance*, *2*, 249– 251.

Held, J. D., & Foley, P. P. (1994). Explanations for accuracy of the general multivariate formulas in correcting for range restriction. *Applied Psychological Measurement*, *18*, 355-367.

Hearn, J. C. (1984). The Relative roles of academic, ascribed, and socioeconomic characteristics in college destinations. *Sociology of Education*, *57*, 22-30.

Higher Education Research Institute (2001). *An overview of the 2000 freshman norms*. Los Angeles: University of California at Los Angeles.

Hillegas, M. B. (1912). A scale for the measurement of quality of English compositions. *Teachers College Record*.

Hobart, T. (1997). President's perspective. *New York Teacher*.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting. Ecology and Biology of Parenting* (Vol. 2, p. 161–188). Mahwah, N.J.: Lawrence Erlbaum.

Horn, L., & West, J. (1992). A profile of parents of eighth graders. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Houston, W. M., & Novick, M. R. (1987). Race-based differential prediction in Air Force technical training programs. *Journal of Educational Measurements*, *24*, 309 –320.

Jensen, A. R. (1980). *Bias in mental testing.* New York: Free Press.

Kanarek, E. A. (1989). *Exploring the murky world of admissions predictions.* Paper presented at the Annual Forum of the Association for Institutional Research, Baltimore, MD.

Karen, D. (1991). The politics of class, race, and gender: Access to higher education in the United States, 1960-1986. *American Journal of Education, 99*, 208-237.

Kao, G., & Tienda, M., (1998). Educational aspirations of minority youth. *American Journal of Education, 106*, 349–384.

King, T., & Bannon, E. (2002). *At what cost? The price that working students pay for a college education*. (PIRGs' Higher Education Project). Washington, DC: State PIRGs.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average.* (College Board Research Report No. 2008-5). New York: The College Board.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate student success. *Science*, *315*, 1080-1081.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*, 162-181.

Kuncel, N. R., & Sackett, P. R. (2007). Selective citation mars conclusions about test validity and predictive bias. *American Psychologist, 61*, 145-146.

LaHuis, D.M., Martin, N.R. and Avis, J.M. (2005) Investigating nonlinear conscientiousness–job performance relations for clerical employees. Human Performance, 18, 199–212.

Lareau, A. (1987). Social class differences in family-school relationships: The importance of cultural capital. *Sociology of Education. 60,* 73–85.

Ledvinka, J. (1979). The statistical definition of fairness in the federal selection guidelines and its implications for minority employment. *Personnel Psychology, 32*, 551-562.

Macauley, F. R. (1931). *The smoothing of time series*. New York: National Bureau of Economic Research.

Matthews, W. (1984). Influences on the learning and participation of minorities in mathematics. *Journal for Research in Mathematics Education, 15,* 84-95.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlations coefficients. *Psychological Methods, 1*, 30-46.

McSpirit, S., & Jones, K.E. (1999). Grade inflation rates among different ability students, controlling for other factors. *Education Policy Analysis Archives, 7*, 30.

Morgan, R. (1989). *Analysis of the predictive validity of the SAT and high school grades from 1976 to 1983* (College Board Rep. No. 89–7). New York: College Board.

Mueller, C.W., & Parcel, T.L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development, 52*, 13–30.

Munday, L. A. (1970). Factors influencing the predictability of college grades. *American Educational Research Journal*, *7,* 99-108.

National Association of Colleges and Employers. (2005). Job Outlook 2005. Retrieved from http://www.unl.edu/svcaa/documents/how_employers_see_candidates.pdf.

National Center for Education Statistics (NCES) (1984). High school course grade standards. *NCES Bulletin* (Report No. NCES-221b). Washington, DC: National Center for Education Statistics.

National Center for Education Statistics. (2009). America's high school graduates: Results of the 2009 NAEP high school transcript study. Retrieved from http://nces.ed.gov/

National Center for Education Statistics. (2011). The condition of education 2011. Retrieved from http://nces.ed.gov/

Nezlek, J. B., & Zyzniewski, L. E. (1998). Using hierarchical linear modeling to analyze grouped data. *Group Dynamics: Theory, Research, and Practice*, *2*, 313-320.

Noble, J., & Sawyer, R. (2002). Predicting different levels of academic success in college using high school GPA and ACT composite score. ACT Research Report Series #2002-4. Iowa City, IA: ACT.

O'Connor, B. P. (2005). Graphical analyses of personality disorders in five-factor model space. *European Journal of Personality*, *19*, 287–305.

Orfield, G., Eaton, S. E., Jones, E.R. (1996). *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education*. New York: New Press.

Park, G., Lubinski, D., Benbow, C. P. (2007). Contrasting intellectual patterns predict creativity in the arts and sciences: Tracking intellectually precocious youth over 25 years. *Psychological Science*, *18*, 948-953.

Partonen, T., Haukka, J., Viilo, K., Hakko, H., Pirkola, S., Isometsa, E., et al. (2004). Cyclic time patterns of death from suicide in northern Finland. *Journal of Affective Disorders*, *78*, 11-19.

Powers, D. E. (2004). Validity of Graduate Record Examinations (GRE) general test scores for admissions to colleges of veterinary medicine. *Journal of Applied Psychology*, *89*, 208-219.

Ramist, L. (1984). Validity of the ATP tests. In T. Donlon (Ed.), *The College Board technical handbook for the Scholastic Aptitude Test and Achievement Tests* (pp. 141–170). NewYork: College Board.

Ramist, L., Lewis, C., & McCamley, L. (1990). Implications of using freshman GPA as the criterion for the predictive validity of the SAT. In Willingham, W.W., Lewis, C., Morgan, R., & Ramist, L., *Predicting college grades: An analysis of* *institutional trends over two decades*, 253-88. Princeton, N.J.: Educational Testing Service.

Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex, language, and ethnic groups* (Research report #93-1). New York: College Board.

Ramist, L., and Weiss, G. (1990). The predictive validity of the SAT, 1964 to 1988. In W. Willingham, C. Lewis, R. Morgan, and L. Ramist (Eds.), *Predicting college grades: An analysis of institutional trends over two decades.*

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297-330.

Roth, P. L., BeVier, C. A., Switzer, F. S. III., & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology, 81,* 548-556.

Ruben, R. J. (1997). A Time Frame of Critical/Sensitive Periods of Language Development, *Acta Otolaryngol*, 117, 202-205.

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, *107*, 1999–2045.

Sackett, P. R., & Wilk, S. L. (1994). Within-Group norming and other forms of score adjustment in preemployment testing. *American Psychologist, 49*, 929-954.

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admission tests and post-secondary academic performance? *Psychological Bulletin*, *135*, 1-22.

Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2011). *The role of socio-economic status in college admissions: Critical tests of claims that admission test validity is artifactual*. Manuscript submitted for publication.

Sackett, P. R., & Yang, H. (2000). Correcting for range restriction: An expanded typology. *Journal of Applied Psychology*, *85*, 112-118.

Schmidt, F. L. (1988). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior, 33,* 272–292.

Schmidt, F. L., Pearlman, K., & Hunter, J. E. (1980). The validity and fairness of employment and educational tests for Hispanic Americans: A review and analysis. *Personnel Psychology*, *33*, 705-724.

Shen, W., Kiger, T. B., Sackett, P. R., Kuncel, N. R., Walmsley, P. T., Beatty, A. S., & Rigdon, J. L. (2011). *Differential prediction in the use of the SAT and high school grades in predicting college performance: Joint effects of race and language*. Manuscript submitted for publication.

Shen, W., Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., & Kiger, T. B. (in press). All validities are not created equal: Determinants of variation in SAT validity across schools. *Applied Measurement in Education*.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*, 417–453.

Scocca, T. (1998, April 23). The great grade-inflation lie. *The Boston Phoenix*.

Stanley, J. C., Benbow, C. P., Brody, L. E., Dauber, S., and Lupowski, A. E. (1992). Gender differences on eighty-six nationally standardized aptitude and achievement tests. In N. Colangelo, S. Assouline, and D. Ambroson (Eds.) *Talent development: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 41–48). Unionville, NY: Trillium Press.

Stern, D., & Nakat, Y. (1991). Paid employment among U.S. college students: Trends, effects, and possible outcomes. *The Journal of Higher Education, 62*, 25-43.

Stumpf, H., & Stanley, J. C. (2002). Group data on high school grade point averages and scores on academic aptitude tests as predictors of institutional graduation rates. *Educational and Psychological Measurement, 62*, 1042-1052.

Stricker, L. J. (2004). The performance of native speakers of English and ESL speakers on the computer-based TOEFL and GRE general test. *Language Testing*, *21*, 146-173.

Talento-Miller, E., & Rudner, L. M. (2005). *Which programs have the highest validity: Identifying characteristics that affect prediction of success*. Research Report #05-03. McLean, VA: Graduate Management Admission Council.

Tate, W. F. (1997). Race-ethnicity, SES, gender, and language proficiency trends in mathematics achievement: An update. *Journal for Research in Mathematics Education, 28,* 652–679.

University of California (1999). *Admissions as a Freshman.* http://www.ucop.edu/pathways/infoctr/introuc/fresh.html

Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcome. *The Review of Higher Education, 27,* 45-73.

Willingham, W. W., & Cole, N. S. (1997). Gender and fair assessment. Mahwah, NJ: Lawrence Erlbaum Associates.

Willingham, W. W. (1985). Success in college: The role of personal qualities and academic ability. New York: College Entrance Examination Board.

Wilson, K. M. (1983). A review of research on the prediction of academic performance after the freshman year (Research report #83-2). New York: College Board: College Entrance Examination Board.

Yamasaki, A., Chinami, M., Morgenthaler, S., Kaneko, Y., Nakashima, K., & Shirakawa, T. (2004). Enterprise failures correlate positively with suicide rate for both sexes in Japan. *Psychological Reports, 95*, 917-920.

Young, J. W., & Johnson, P. M. (2004). The impact of an SES-based model on a college's undergraduate admissions outcomes. *Research in Higher Education, 45,* 777–797.

Zak, P. M., Benbow, C. P., & Stanley, J. C. (1983). Several factors associated with success as an undergraduate chemistry major. *College and University*, *58*, 303-312.

Zwick, R., & Schlemer, L. (2004). SAT validity for linguistic minorities at the University of California, Santa Barbara. *Educational Measurement: Issues and Practice*, *2*3, 6-16.

Zwick, R., & Sklar, J. C. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal*, *42*, 439-464.

Appendix A: Tables

Table 1: *Description of variables included in each school characteristic factor/cluster*

Size & Financial Need		Selectivity	Cost	Homogeneous campus life	Math and sci. prep.	Classic predictors	Alternative predictors
Freshmen class size reporting SAT scores	Number of students that received aid	Mean freshmen GPA	Cost of tuition/fees	Public vs. private	Number of science units required	Importance of school record	Importance of recommend-dations
Total number of undergraduate students	Number of undergraduate students awarded non-need based aid	Freshmen GPA standard deviation	Cost of out-of-state tuition/fees	Percentage of students living on campus	Number of math units required	Importance of test scores	Importance of extracurricular activities
Total number of graduate students	Number of freshmen awarded non-need based aid	Mean SAT composite score (CR+M+W)	Cost of room/board	Pct. of part time students			Importance of essays
Number of full-time faculty	Number of students that applied for financial aid	Percentage of students admitted		Percentage of out of state students			
Percentage of classes with 50-99 students	Number of students judged to have aid	Pct. of students with HSGPA of 3.5 or higher					
Percentage of classes with 100+ students		Pct. of students ranked in the top quarter of high school					
		Pct. of students ranked in the top ten pct. of high school					
		Percentage of students receiving scholarships					
		Average scholarship amount awarded					

Table 2

Correlations among institutional factors and HSGPA-FGPA validity

	1	2	3	4	5	6	7	8	9	10	11
1. selectivity factor											
2. size & financial need factor	-0.03										
3.cost factor	0.71**	-0.30**									
4. homogeneous campus life factor	0.71**	-0.33**	0.72**								
5.classic predictors factor	-0.06	0.22*	0.04	-0.07							
6. alternative predictors factor	0.14	-0.17	0.30**	0.18	0.51**						
7. math & science prep factor	0.26**	-0.02	0.19*	0.24*	0.21*	0.12					
8. second order composite	0.90**	-0.24*	0.90**	0.90**	-0.03	0.23*	0.25**				
9. percent disadvantaged minorities	-0.15	0.80	-0.13	-0.21*	-0.09	-0.07	-0.07	-0.18			
10. percent female	-0.23*	-0.30**	-0.06	-0.05	-0.15	-0.072	-0.02	-0.12	0.31**		
11. School-specific validity	-0.17	-0.17	-0.10	0.08	-0.001	-0.129	-0.03	-0.07	-0.23*	0.09	
12. National-population validity	0.17	-0.27**	0.17	0.31**	-0.05	-0.126	0.06	0.24*	-0.31**	0.07	0.84**

Notes. N=110 schools. *p<.05 **p<.01

Table 3

Institutional predictors predicting school specific range restriction and unreliability corrected HSGPA validity

	Unstandardized B	Standardized Beta	Sig.
Constant	**0.50**		**0.00**
Selectivity factor	-0.02	-0.28	0.08
Size and financial need factor	-0.01	-0.14	0.22
Cost factor	-0.01	-0.17	0.28
Homogeneous campus life factor	**0.03**	**0.36**	**0.02**
Classic predictors factor	0.01	0.15	0.21
Alternative predictors factor	-0.02	-0.21	0.07
Math and science preparation factor	0.00	-0.03	0.74
Percentage of disadvantaged minorities	**0.001**	**-0.23**	**0.02**
Percentage of female students	0.00	0.06	0.54
R^2	.20		
Adjusted R^2	.13		

Note. N=110 schools.

Table 4

Second order selectivity composite and other institutional predictors predicting school specific range restriction and unreliability corrected HSGPA validity

	Unstandardized B	Standardized Beta	Sig.
Constant	**0.49**		**0.00**
Second order composite (selectivity, cost, homogeneity)	0.00	-0.09	0.37
Size and financial need factor	**-0.02**	**-0.22**	**0.04**
Classic predictors factor	0.01	0.17	0.17
Alternative predictors factor	**-0.02**	**-0.24**	**0.04**
Math and science preparation factor	0.00	-0.03	0.74
Percentage of disadvantaged minorities	**-0.002**	**-0.26**	**0.01**
Percentage of female students	0.00	0.09	0.37
R^2	.15		
Adjusted R^2	.09		

Note. N=110 schools.

Table 5

Institutional predictors predicting national range restriction and unreliability corrected HSGPA validity

	Unstandardized B	Standardized Beta	Sig.
Constant	**0.57**		**0.00**
Selectivity factor	0.01	0.11	0.49
Size and financial need factor	**-0.02**	**-0.26**	**0.02**
Cost factor	-0.01	-0.09	0.54
Homogeneous campus life factor	0.02	0.24	0.11
Classic predictors factor	0.01	0.18	0.12
Alternative predictors factor	**-0.02**	**-0.30**	**0.01**
Math and science preparation factor	0.00	-0.03	0.77
Percentage of disadvantaged minorities	**-0.001**	**-0.27**	**0.01**
Percentage of female students	0.00	0.11	0.28
R^2	.26		
Adjusted R^2	.19		

Note. N=110 schools.

Table 6

Second order selectivity composite and other institutional predictors predicting national range restriction and unreliability corrected HSGPA validity

	Unstandardized B	Standardized Beta	Sig.
Constant	**0.58**		**0.00**
Second order composite (selectivity, cost, homogeneity)	**0.01**	**0.22**	**0.02**
Size and financial need factor	**-0.02**	**-0.25**	**0.02**
Classic predictors factor	0.01	0.16	0.15
Alternative predictors factor	**-0.02**	**-0.31**	**0.01**
Math and science preparation factor	0.00	-0.01	0.88
Percentage of disadvantaged minorities	**-0.001**	**-0.29**	**0.00**
Percentage of female students	0.00	0.11	0.27
R^2	.25		
Adjusted R^2	.20		

Note. N=110 schools.

Table 7

HLM results for institutional characteristics predicting HSGPA slope

	Coefficient	Std. Error	Sig.
For Intercept			
Intercept	**3.02**	**0.02**	**0.00**
For HSGPA Slope			
Intercept	**0.54**	**0.01**	**0.00**
Selectivity factor	**-0.06**	**0.01**	**0.00**
Size and financial need factor	-0.02	0.01	0.07
Cost factor	**-0.05**	**0.02**	**0.00**
Homogeneous campus life factor	0.03	0.02	0.09
Classic predictors factor	0.02	0.01	0.06
Alternative predictors factor	-0.03	0.01	0.06
Math and science preparation factor	0.00	0.01	0.95
Percentage of disadvantaged minorities	0.00	0.00	0.13
Percentage of female students	0.00	0.00	0.11
Sigma Squared	0.39		
Tau	.03		
Random effect	Std. Dev.	Variance	Sig.
Intercept 1	0.18	0.03	0.00
Level 1	0.62	0.39	

Table 8

HLM results for second order composite and other institutional characteristics predicting HSGPA slope

	Coefficient	Std. Error	Sig.
For Intercept			
Intercept	**3.01**	**.02**	**0.00**
For Slope			
Intercept	**0.54**	**0.01**	**0.00**
Second order composite (selectivity, cost, homogeneity)	**-0.03**	**0.00**	**0.00**
Size and financial need factor	**-0.03**	**0.01**	**0.01**
Classic predictors factor	0.01	0.01	0.327
Alternative predictors factor	-0.03	0.02	0.064
Math and science preparation factor	0.00	0.01	0.90
Percentage of disadvantaged minorities	**0.00**	**0.00**	**0.03**
Percentage of female student	0.00	0.00	0.10
Sigma Squared	0.39		
Tau	.03		

Random effect	Std. Dev.	Variance	Sig.
Intercept 1	0.18	0.03	0.00
Level 1	0.62	0.39	

Table 9

Overall descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
HSGPA	61675	1.20	5.00	3.53	0.44
FGPA	61675	-4.86	1.70	-0.51	1.06
Humanities GPA	58108	-6.76	1.64	-0.43	1.05
STEM GPA	55895	-4.90	1.50	-0.42	1.04
Social Sciences GPA	51165	-5.02	1.49	-0.48	1.05

Table 10

Correlational matrices for HSGPA, college GPA and adjustments to college GPA

	HSGPA	FGPA	Humanities GPA	STEM GPA	Social Sciences GPA
HSGPA	1	0.53	0.48	0.45	0.46
FGPA	0.53	1	0.82	0.83	0.79
Humanities GPA	0.48	0.82	1	0.57	0.59
STEM GPA	0.45	0.83	0.57	1	0.57
Social Sciences GPA	0.46	0.79	0.59	0.57	1
Adjustments to college GPA					
FGPA		1	0.999	0.998	0.997
Humanities GPA		0.999	1	0.998	0.997
STEM GPA		0.998	0.998	1	0.995
Social Sciences GPA		0.997	0.997	0.997	1

Table 11

Results of power polynomial analyses for the full dataset and for those scoring 1 SD above the mean

	Full Data				Top End			
Criteria	R^2 linear	R^2 quadratic	Difference	Sig. of Difference	R^2 linear	R^2 quadratic	Difference	Sig. of Difference
FGPA	.283	.285	.002	.000	.013	.015	.002	.000
Humanities GPA	.228	.229	.001	.000	.009	.010	.001	.000
STEM GPA	.204	.208	.004	.000	.012	.013	.001	.000
Social Sciences GPA	.207	.209	.002	.000	.007	.010	.003	.000

Table 12

Descriptive statistics by ethnicity

	N	Minimum	Maximum	Mean	Std. Deviation
White					
HSGPA	42737	1.20	4.99	3.53	0.45
FGPA	42737	-4.86	1.70	-0.46	1.05
Humanities GPA	40232	-6.76	1.64	-0.39	1.02
STEM GPA	38583	-4.90	1.50	-0.37	1.03
Social Sciences GPA	35412	-5.01	1.49	-0.43	1.03
Asian					
HSGPA	7044	1.70	4.95	3.61	0.39
FGPA	7044	-4.53	1.70	-0.46	1.06
Humanities GPA	6583	-6.39	1.64	-0.36	1.06
STEM GPA	6593	-4.90	1.41	-0.34	1.02
Social Sciences GPA	5844	-5.02	1.49	-0.46	1.07
Black					
HSGPA	3455	1.30	4.73	3.35	0.47
FGPA	3455	-4.76	1.52	-1.06	1.06
Humanities GPA	3301	-6.39	1.44	-0.87	1.10
STEM GPA	3195	-4.24	1.41	-0.93	1.07
Social Sciences GPA	2912	-4.94	1.26	-0.87	1.11
Hispanic					
HSGPA	3147	1.67	4.90	3.49	0.43
FGPA	3147	-4.74	1.52	-0.79	1.06
Humanities GPA	2991	-6.39	1.44	-0.66	1.09
STEM GPA	2830	-4.19	1.41	-0.67	1.08
Social Sciences GPA	2591	-5.01	1.34	-0.72	1.08

Table 13

Descriptive statistics for language groups

	N	Minimum	Maximum	Mean	Std. Dev.
Asian English Only					
HSGPA	5408	1.70	4.95	3.61	0.39
FGPA	5408	-4.53	1.63	-0.45	1.06
Humanities GPA	5079	-6.39	1.64	-0.35	1.06
STEM GPA	5046	-4.90	1.41	-0.35	1.03
Social Sciences GPA	4519	-5.02	1.49	-0.44	1.06
Asian Bilingual					
HSGPA	1206	2.20	4.83	3.60	0.39
FGPA	1206	-3.86	1.70	-0.50	1.04
Humanities GPA	1106	-6.29	1.58	-0.39	1.06
STEM GPA	1143	-3.82	1.41	-0.37	1.01
Social Sciences GPA	974	-5.02	1.49	-0.51	1.09
Asian Non-English					
HSGPA	381	2.13	4.72	3.61	0.36
FGPA	381	-3.06	1.55	-0.44	1.06
Humanities GPA	350	-4.48	1.44	-0.44	1.06
STEM GPA	358	-3.58	1.33	-0.23	0.98
Social Sciences GPA	310	-4.95	1.25	-0.59	1.19
Hispanic English Only					
HSGPA	2319	1.67	4.90	3.47	0.43
FGPA	2319	-4.74	1.52	-0.78	1.06
Humanities GPA	2205	-6.39	1.44	-0.66	1.10
STEM GPA	2087	-4.19	1.41	-0.67	1.07
Social Sciences GPA	1902	-5.01	1.34	-0.72	1.07
Hispanic Bilingual					
HSGPA	727	2.13	4.61	3.52	0.41
FGPA	727	-4.26	1.28	-.085	1.07
Humanities GPA	694	-5.06	1.27	-.069	1.08
STEM GPA	654	-4.19	1.33	-.071	1.13
Social Sciences GPA	608	-5.01	1.26	-.073	1.11
Hispanic Non-English					
HSGPA	94	2.24	4.49	3.50	0.45
FGPA	94	-2.76	1.47	-.051	0.95
Humanities GPA	85	-3.00	1.19	-0.50	1.00
STEM GPA	84	-2.97	1.26	-0.28	0.94
Social Sciences GPA	79	-2.98	1.10	-0.45	0.92

Table 14

Descriptive statistics for SES variables by ethnicity

	N	Minimum	Maximum	Mean	Std. Dev.
Overall					
SES	37839	-3.94	1.35	0.00	1.00
Father's Educ.	37839	8	18	15.17	2.47
Mother's Educ.	37839	8	18	14.92	2.35
Income	37839	10000	100000	74812	27412
White					
SES	28900	-3.94	1.35	0.15	0.87
Father's Educ.	28900	8	18	15.39	2.31
Mother's Educ.	28900	8	18	15.15	2.19
Income	28900	10000	100000	79189	24676
Asian					
SES	4417	-3.94	1.35	-0.34	1.23
Father's Educ.	4417	8	18	14.91	2.79
Mother's Educ.	4417	8	18	14.34	2.73
Income	4417	10000	100000	62842	31198
Black					
SES	2281	-3.94	1.35	-0.54	1.08
Father's Educ.	2281	8	18	14.16	2.497
Mother's Educ.	2281	8	18	14.44	2.369
Income	2281	10000	100000	56209	30022
Hispanic					
SES	2241	-3.94	1.35	-0.65	1.24
Father's Educ.	2241	8	18	13.86	2.99
Mother's Educ.	2241	8	18	13.70	2.76
Income	2241	10000	100000	60898	30424

Table 15

SES descriptive statistics for the national population of SAT-takers

	N	Mean	Std. Dev.
Overall			
Father's Educ.	1154376	14.44	2.65
Mother's Educ.	1181978	14.26	2.49
Income	870015	63243	30967
White			
Father's Educ.	716668	14.81	2.41
Mother's Educ.	725225	14.61	2.25
Income	532295	72077	27243
Asian			
Father's Educ.	103779	14.76	2.88
Mother's Educ.	104637	14.16	2.82
Income	77026	53655	32933
Black			
Father's Educ.	119508	13.46	2.37
Mother's Educ.	129782	13.75	2.30
Income	97968	43947	29124
Hispanic			
Father's Educ.	125327	12.77	3.08
Mother's Educ.	130856	12.72	2.89
Income	103344	44702	29823

Table 16

Correlational matrix between SES and key variables

	1	2	3	4	5	6	7	8
1. SES Composite								
2. ln Income	0.75							
3. Father's Education	0.83	0.42						
4. Mother's Education	0.81	0.37	0.55					
5. HSGPA	0.07	0.02	0.09	0.07				
6. FGPA	0.19	0.11	0.15	0.15	0.53			
7. Humanities GPA	0.15	0.09	0.15	0.12	0.48	0.82		
8. STEM GPA	0.15	0.09	0.15	0.12	0.45	0.82	0.57	
9. Social Sciences GPA	0.16	0.09	0.15	0.13	0.46	0.79	0.60	0.58

Notes. Total Sample size is 40,527. Sample sizes for the specific course type GPAs vary from 38,232 to 30,665.

Table 17

Correlation between HSGPA and FGPA by SES group

SES Group	HSGPA-FGPA *r*	n
Bottom 20%	0.47	8077
21%-40%	0.53	8245
41%-60%	0.54	7764
61%-80%	0.56	8198
Top 20%	0.55	8243

Table 18

Intended major and degree goal by SES group

Intended Major	< 21%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
Humanities	10%	12%	14%	14%	16%
Social Sciences	20%	19%	19%	20%	22%
Hard Sciences	44%	41%	41%	38%	36%
Architecture	3%	4%	4%	4%	4%
Business	16%	16%	16%	18%	16%
Education	5%	5%	4%	4%	3%
Technical	0%	0%	0%	0%	0%
Undecided	2%	3%	3%	3%	3%
Degree Goal					
Certification	1%	1%	0%	0%	0%
Associate	0%	0%	0%	0%	0%
Bachelor	22%	26%	26%	24%	13%
Master	32%	33%	35%	35%	36%
Doctoral	26%	20%	21%	22%	30%
Undecided	0%	0%	0%	0%	0%
Other	20%	20%	17%	19%	21%

Table 19

Course-taking patterns by SES group

	0% to 20%			21% to 40%			41% to 60%			61% to 80%			81% to 100%		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
HS English Yrs.	3.84	0.48	7265	3.89	0.44	7363	3.9	0.41	6847	3.91	0.42	7168	3.92	0.42	7224
HS Mathematics Yrs.	3.83	0.55	6767	3.86	0.51	6847	3.90	0.49	6421	3.92	0.48	6743	3.93	0.47	6761
HS Nat. Science Yrs.	3.53	0.74	6676	3.59	0.68	6794	3.63	0.67	6384	3.67	0.66	6714	3.71	0.64	6745
HS Soc. Sci. Yrs.	3.53	0.70	6719	3.61	0.66	6822	3.65	0.64	6415	3.67	0.62	6725	3.71	0.61	6756
HS Pre-Calculus Yrs.	0.96	0.35	4161	0.97	0.33	4226	0.97	0.33	4215	0.98	0.33	4662	0.98	0.31	4903
HS Calculus Yrs.	1.07	0.45	2387	1.08	0.42	2505	1.09	0.44	2738	1.11	0.45	3159	1.08	0.38	3405
HS Physics Yrs.	1.05	0.40	3846	1.08	0.43	3992	1.08	0.41	4048	1.09	0.41	4554	1.09	0.39	4764
Humanities Credits	3.57	2.23	7024	3.68	2.33	7241	3.76	2.42	6849	3.75	2.43	7262	3.86	2.48	7317
STEM Credits	3.52	2.18	6883	3.48	2.21	6994	3.55	2.26	6657	3.55	2.30	6977	3.56	2.36	6887
Soc. Sci. Credits	2.24	1.21	6152	2.25	1.23	6324	2.20	1.20	6121	2.24	1.19	6416	2.24	1.21	6356

Table 20

Percent of students in rigorous coursework and help seeking intentions by SES

	< 21%	21 - 40%	41 - 60%	61 - 80%	81- 100%
HS Honors English	44%	45%	47%	50%	51%
HS Honors Math	36%	37%	39%	43%	44%
HS Honors Nat. Sci.	33%	35%	37%	41%	43%
HS Honors Soc. Sci.	38%	39%	42%	45%	48%
HS Honors Pre-Calc.	18%	19%	21%	24%	25%
HS Honors Calculus	18%	19%	22%	25%	26%
HS Honors Biology	25%	26%	28%	31%	33%
HS Honors Chemistry	24%	24%	27%	30%	31%
HS Honors Physics	16%	17%	19%	22%	23%
College Calculus	26%	25%	29%	31%	32%
College Chemistry	28%	25%	27%	27%	26%
Intentions					
Help w/ Educ. Plans	21%	19%	18%	17%	15%
Help w/ Career Plans	21%	19%	18%	17%	16%
Help w/ Study Skills	27%	25%	23%	21%	19%
Help w/ Math Skills	23%	20%	18%	16%	14%
Help w/ Reading Skills	15%	11%	11%	11%	10%
Help w/ Writing Skills	24%	19%	18%	17%	16%
No Plans for Help	29%	33%	35%	38%	40%
Part-time Job	75%	70%	63%	57%	49%

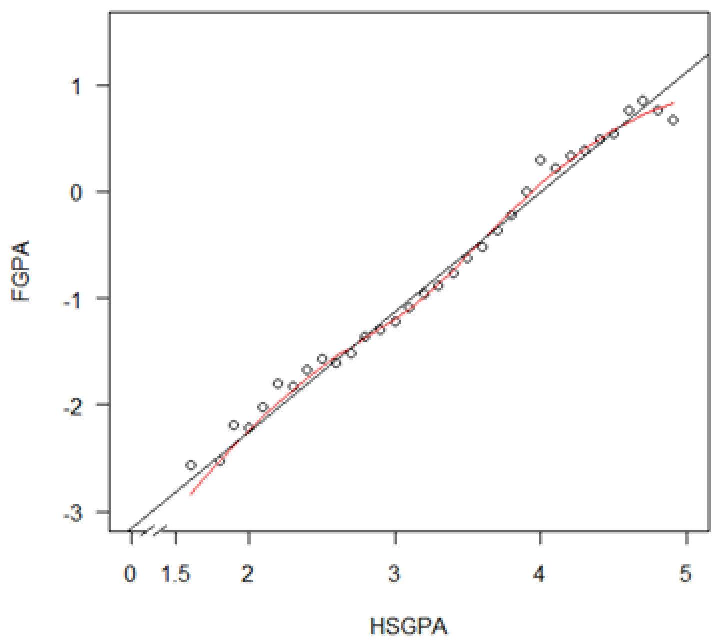
Table 21

Partial Correlations with SES holding SAT constant

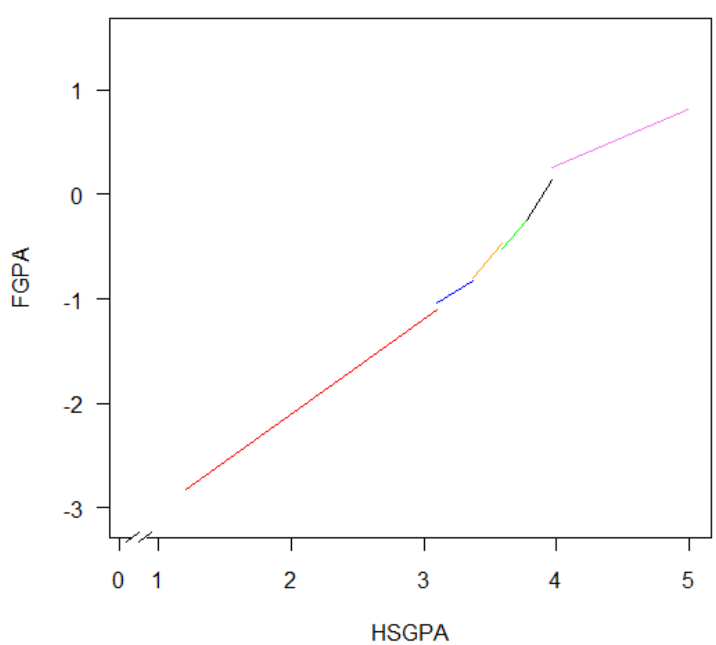
	Partial *r* w/ SES
HS English Yrs.	0.03
HS Mathematics Yrs.	-0.02
HS Nat. Science Yrs.	0.01
HS Soc. Sci. Yrs.	0.04
HS Pre-Calculus Yrs.	-0.02
HS Calculus Yrs.	-0.04
HS Physics Yrs.	-0.01
Humanities Credits	-0.04
STEM Credits	-0.03
Soc. Sci. Credits	0.02
HS Honors English	-0.01
HS Honors Math	0.00
HS Honors Nat. Sci.	-0.05
HS Honors Soc. Sci.	-0.04
HS Honors Pre-Calc.	-0.02
HS Honors Calculus	-0.03
HS Honors Biology	-0.02
HS Honors Chemistry	-0.04
HS Honors Physics	-0.03
College Calculus	-0.04
College Chemistry	-0.03
Help Seeking Comp	-0.06
Part-time Job	-0.13

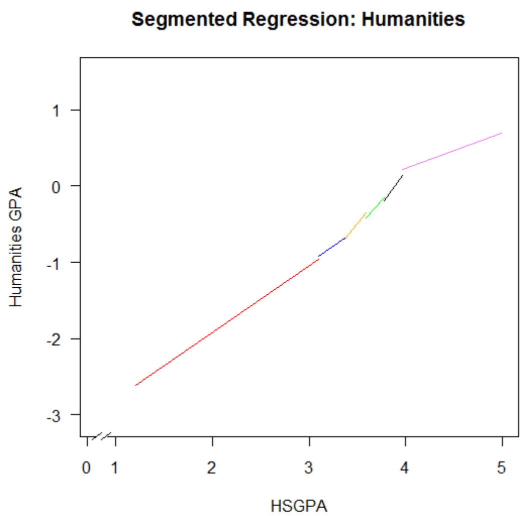
Appendix B: Figures

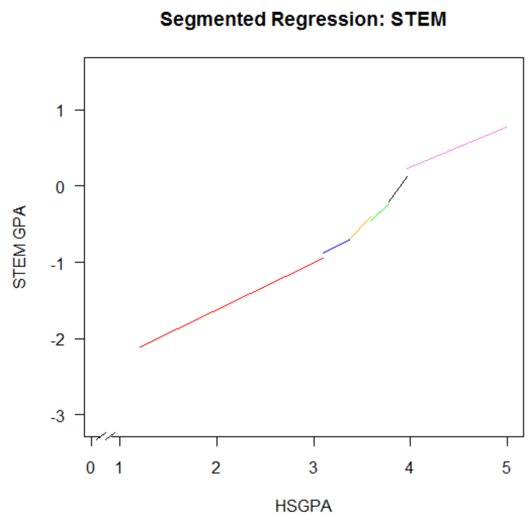
LOESS vs. Linear Regression

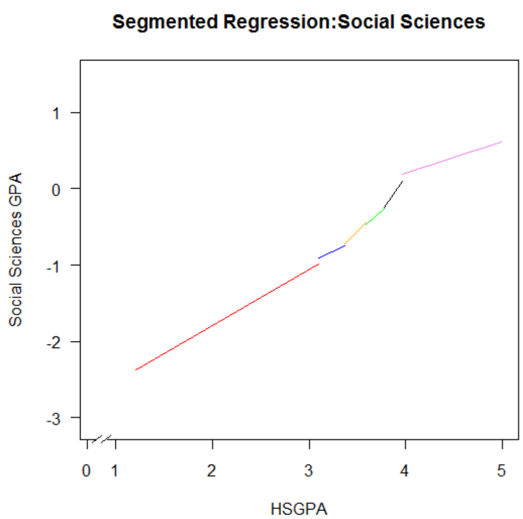


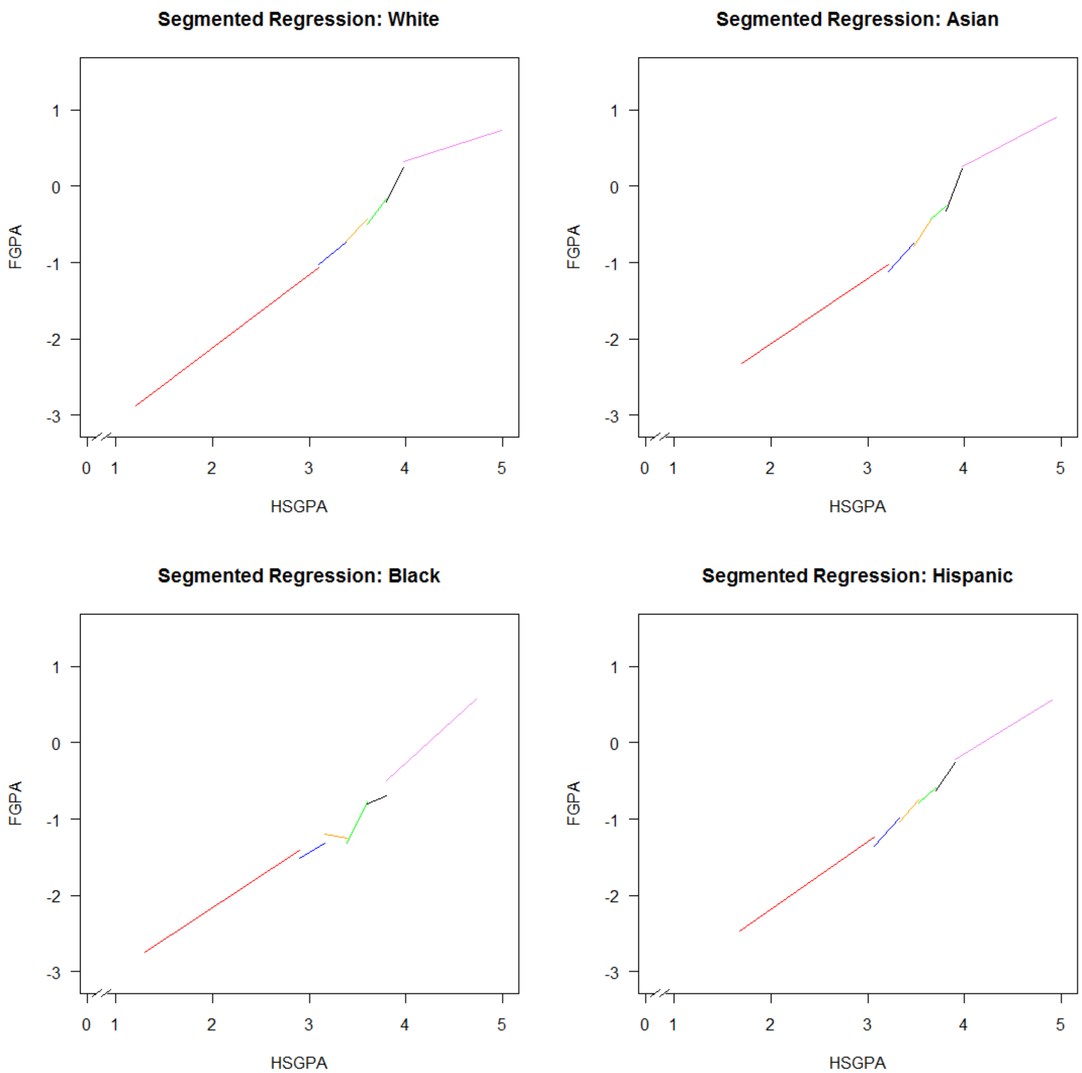
Segmented Regression: FGPA

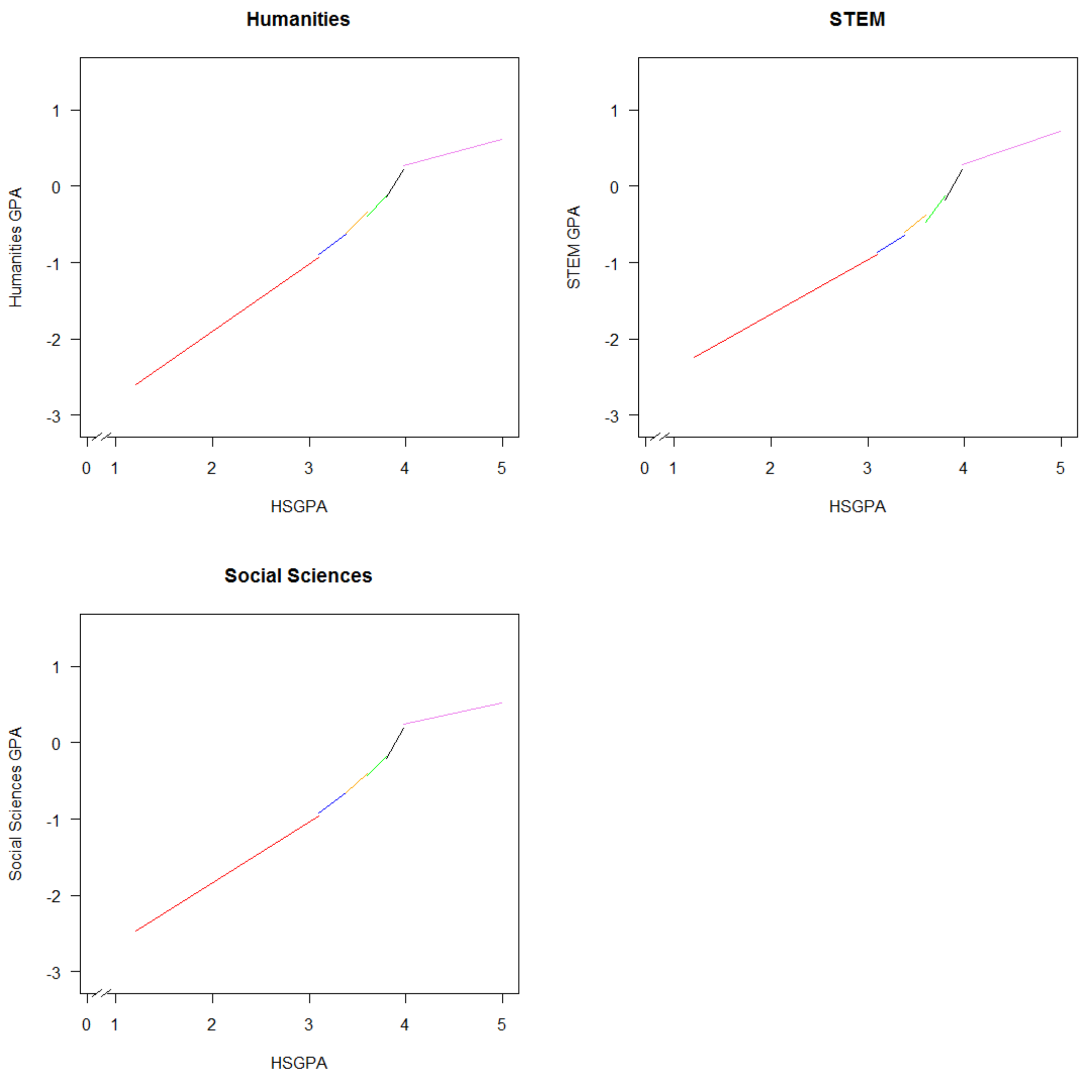


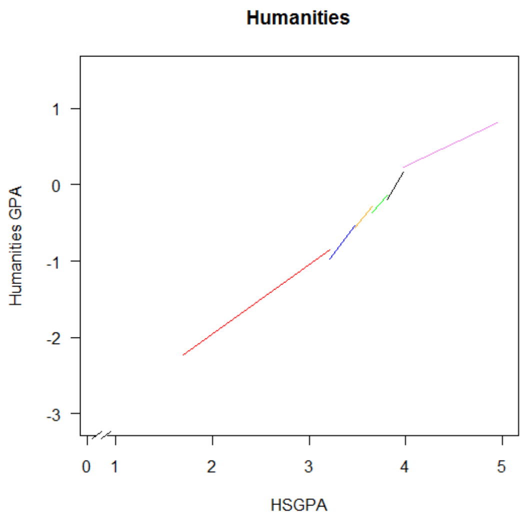

Segmented Regression: Humanities
Humanities GPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5


Segmented Regression: STEM
STEM GPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5

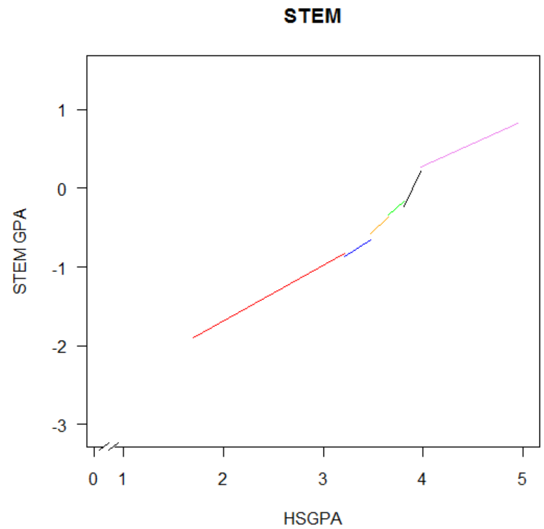

Segmented Regression:Social Sciences
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5


Segmented Regression: White
Segmented Regression: Asian
Segmented Regression: Black
Segmented Regression: Hispanic
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

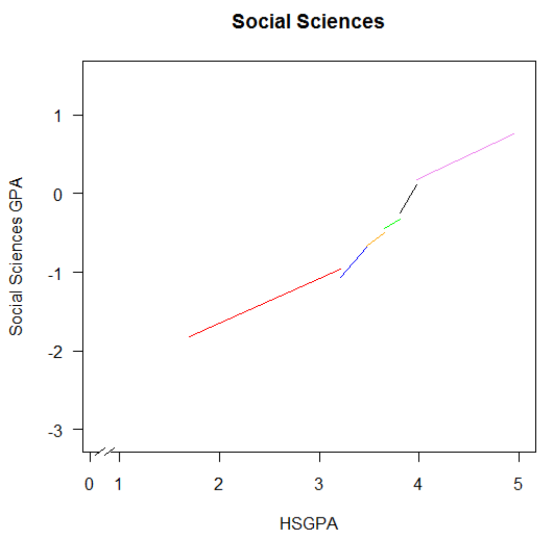

Humanities
Humanities GPA
1
0
-1
-2
-3
0 1 2 3 4 5
HSGPA
STEM
STEM GPA
1
0
-1
-2
-3
0 1 2 3 4 5
HSGPA
Social Sciences
Social Sciences GPA
1
0
-1
-2
-3
0 1 2 3 4 5
HSGPA



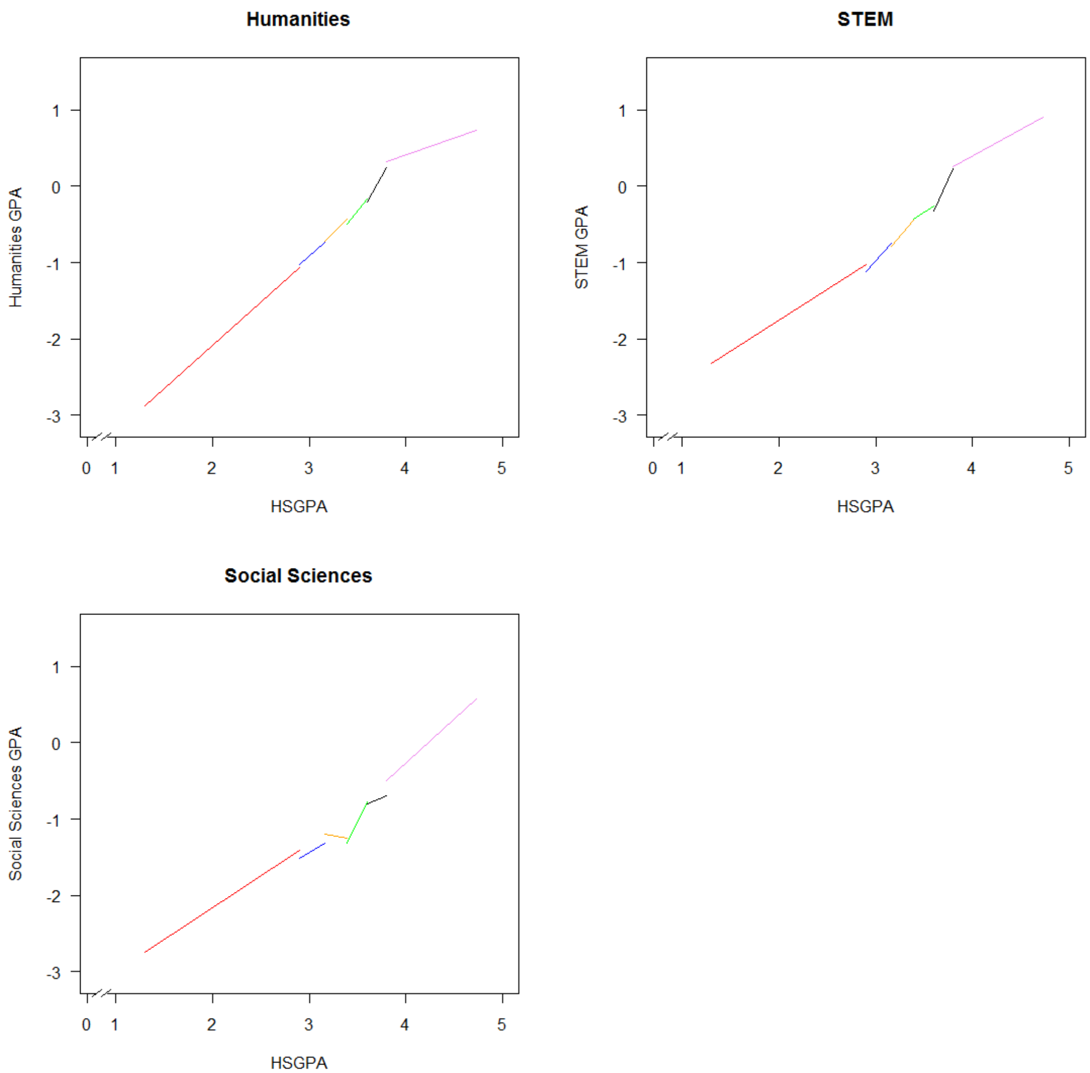
Humanities
Humanities GPA
HSGPA



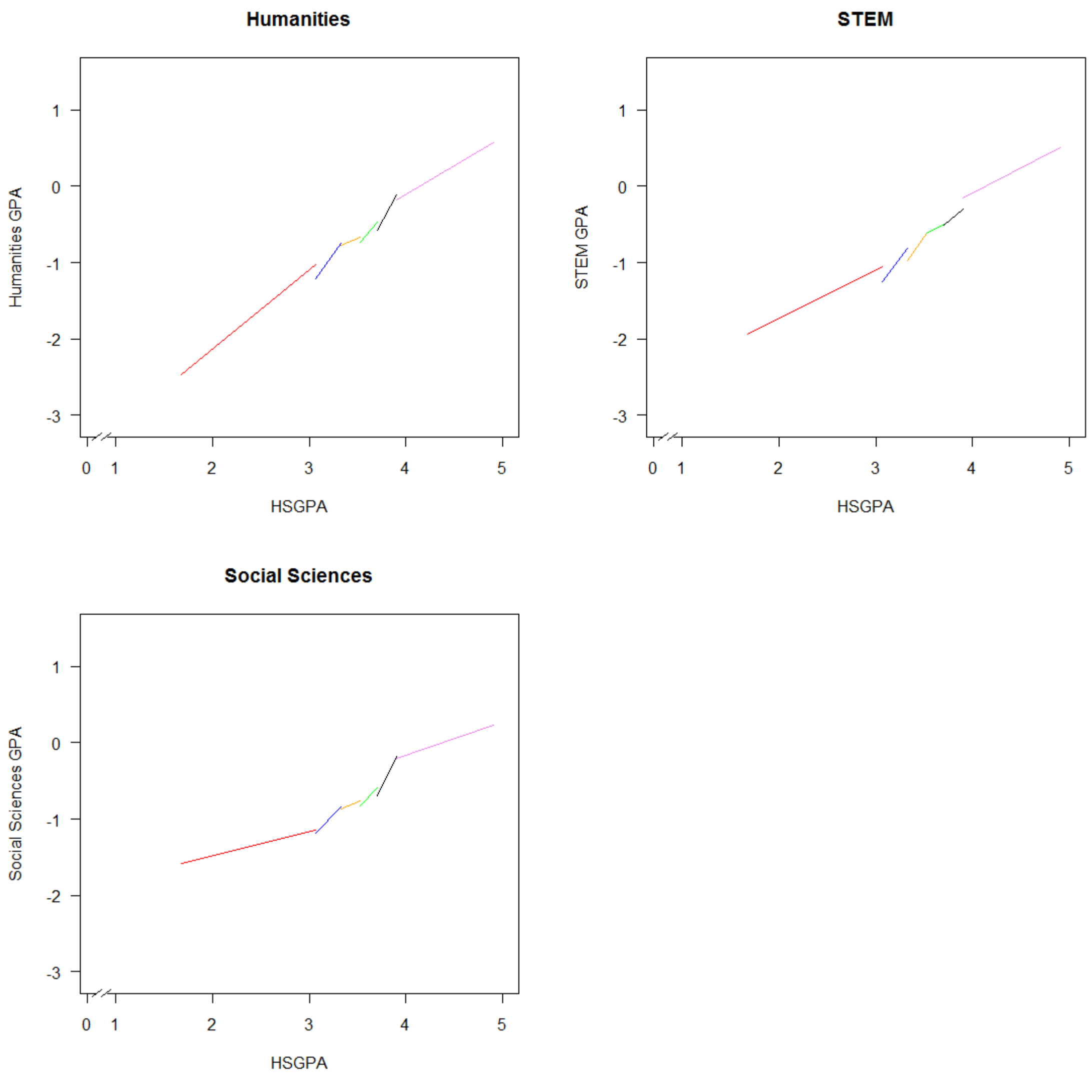
STEM
STEM GPA
HSGPA

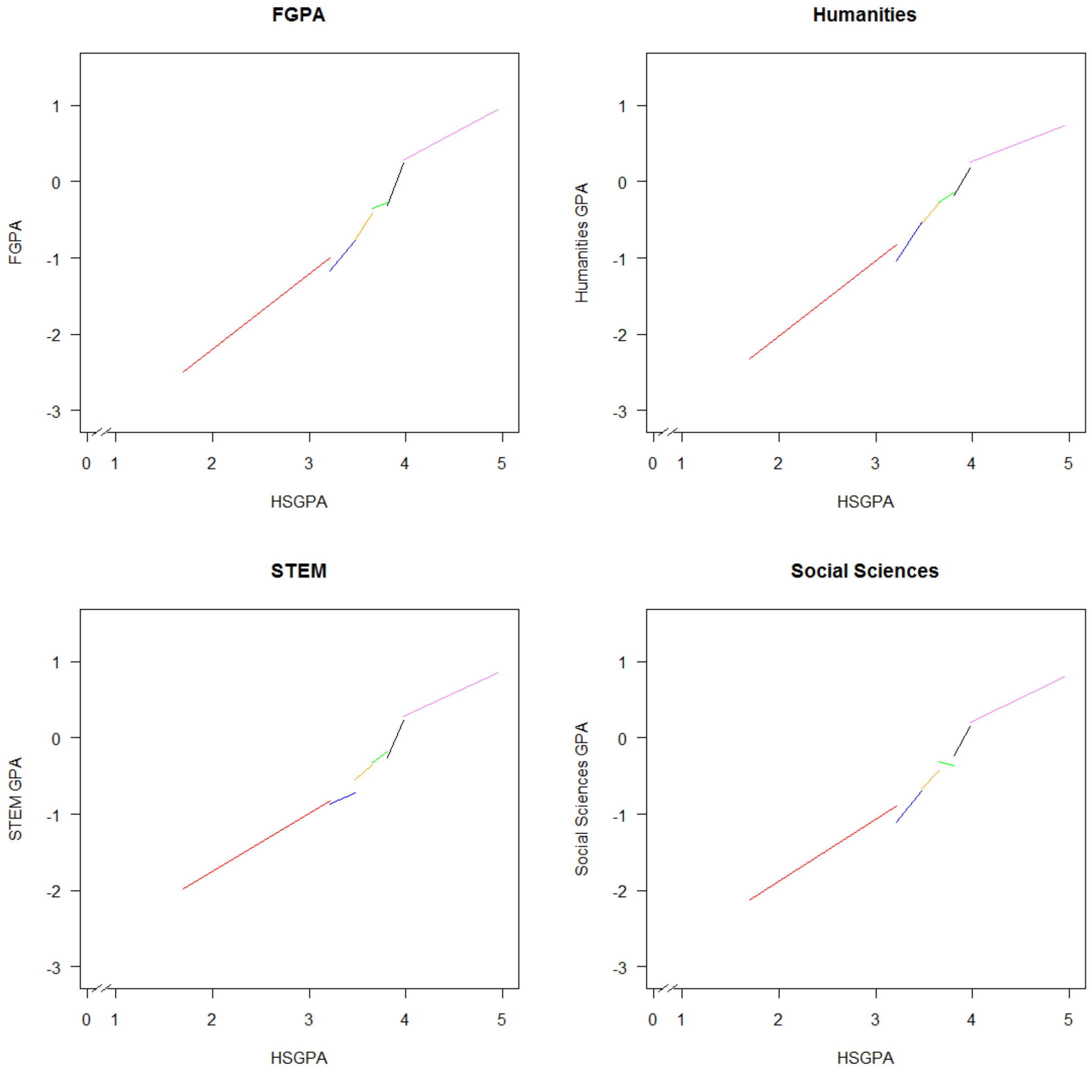


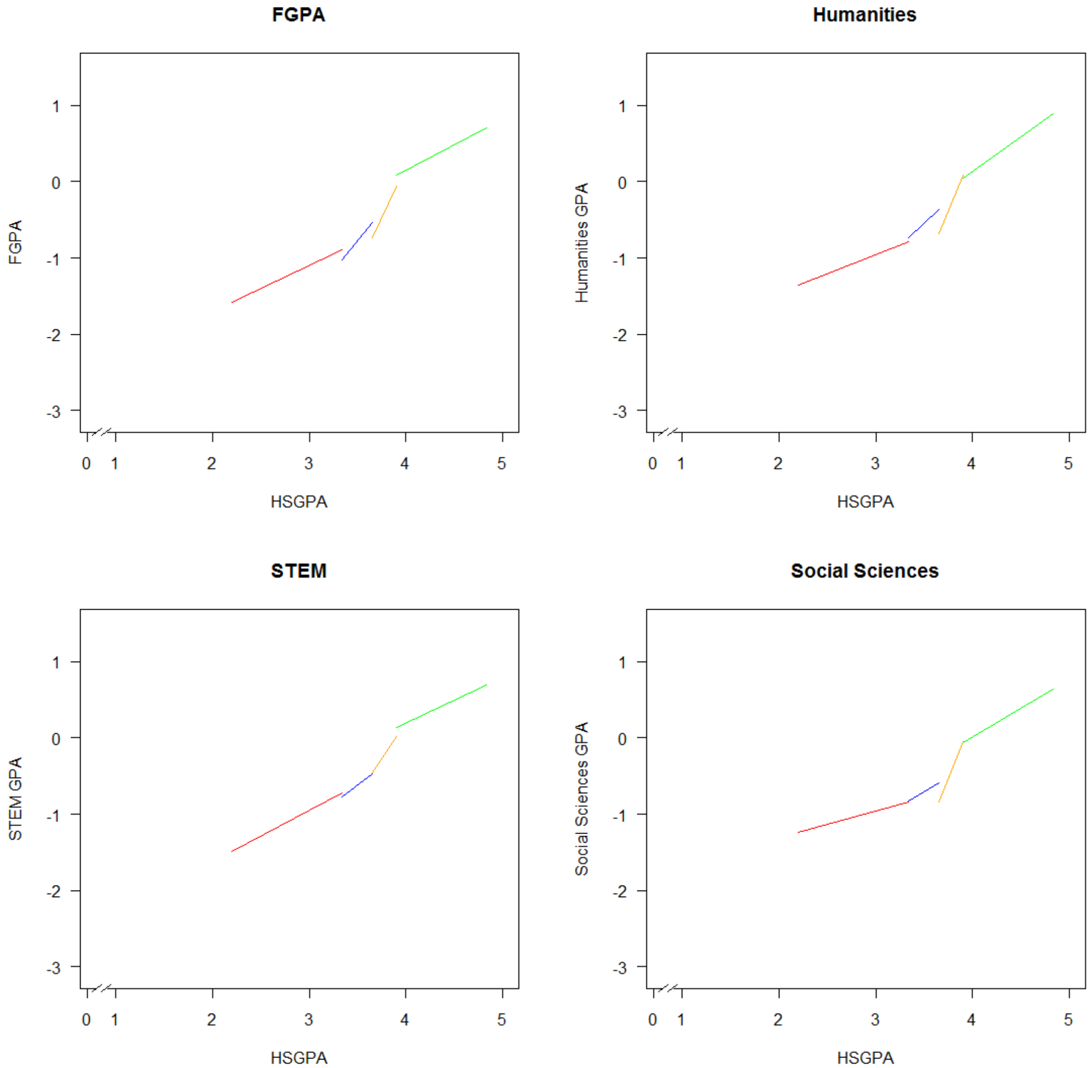
Social Sciences
Social Sciences GPA
HSGPA

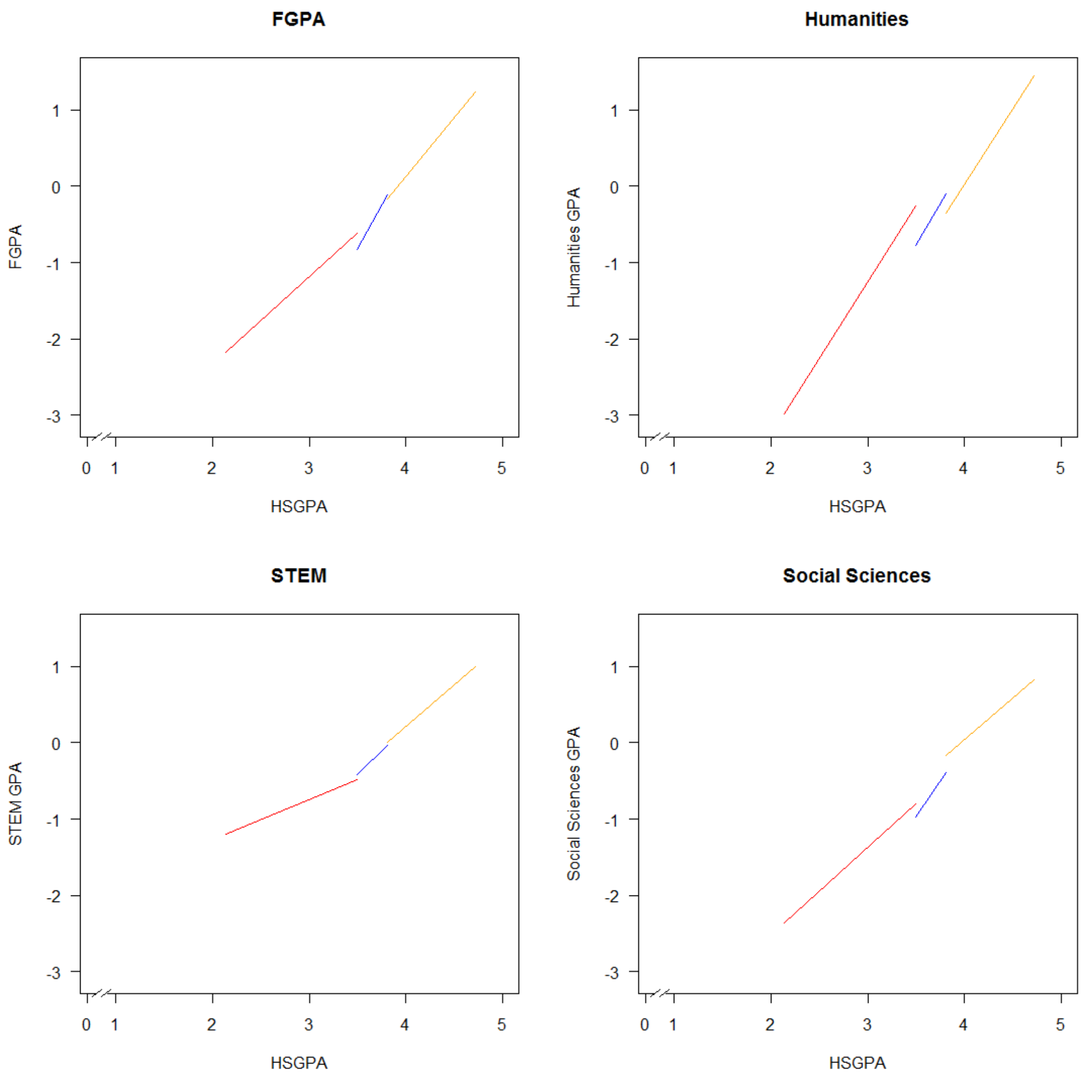


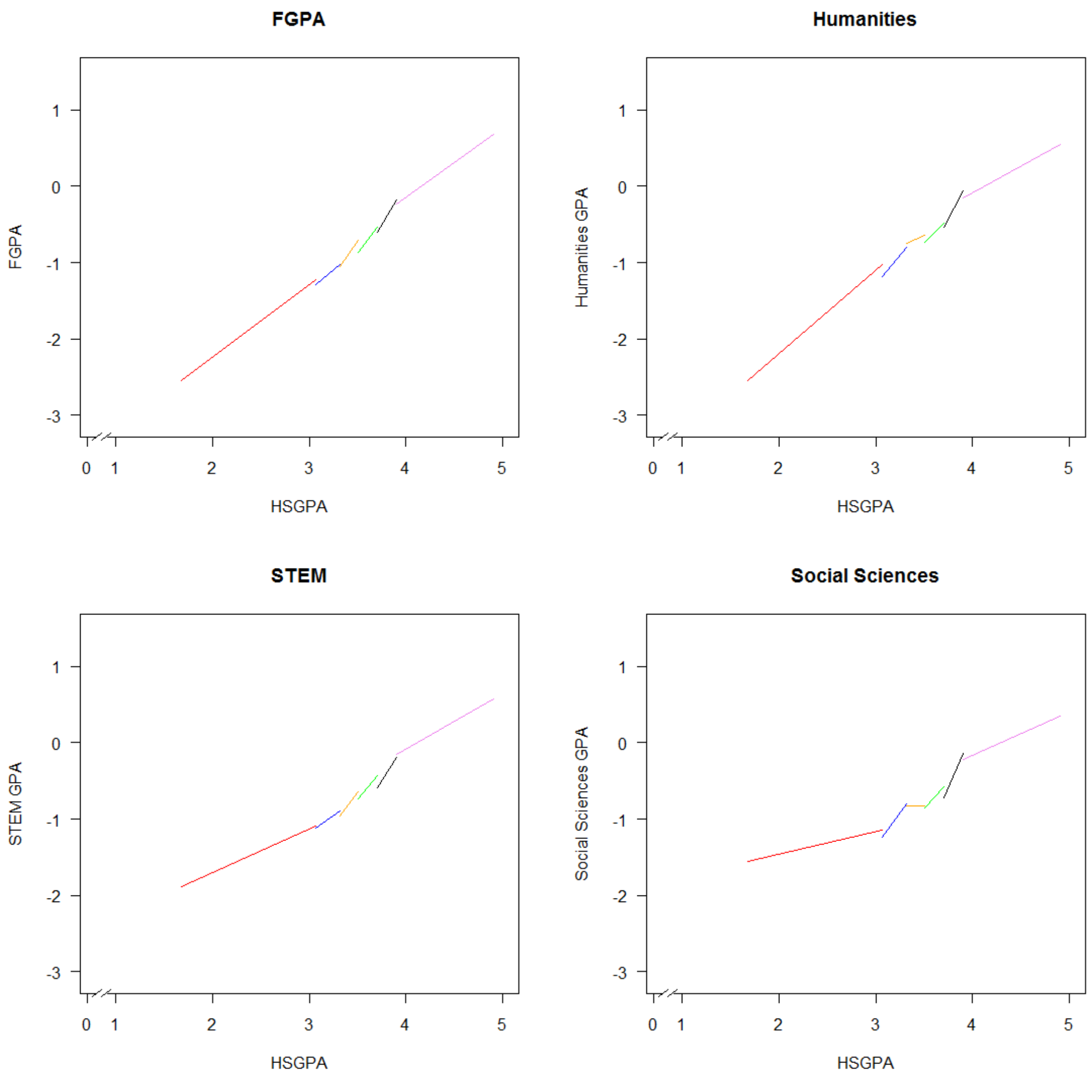
Humanities
Humanities GPA
HSGPA
STEM
STEM GPA
HSGPA
Social Sciences
Social Sciences GPA
HSGPA

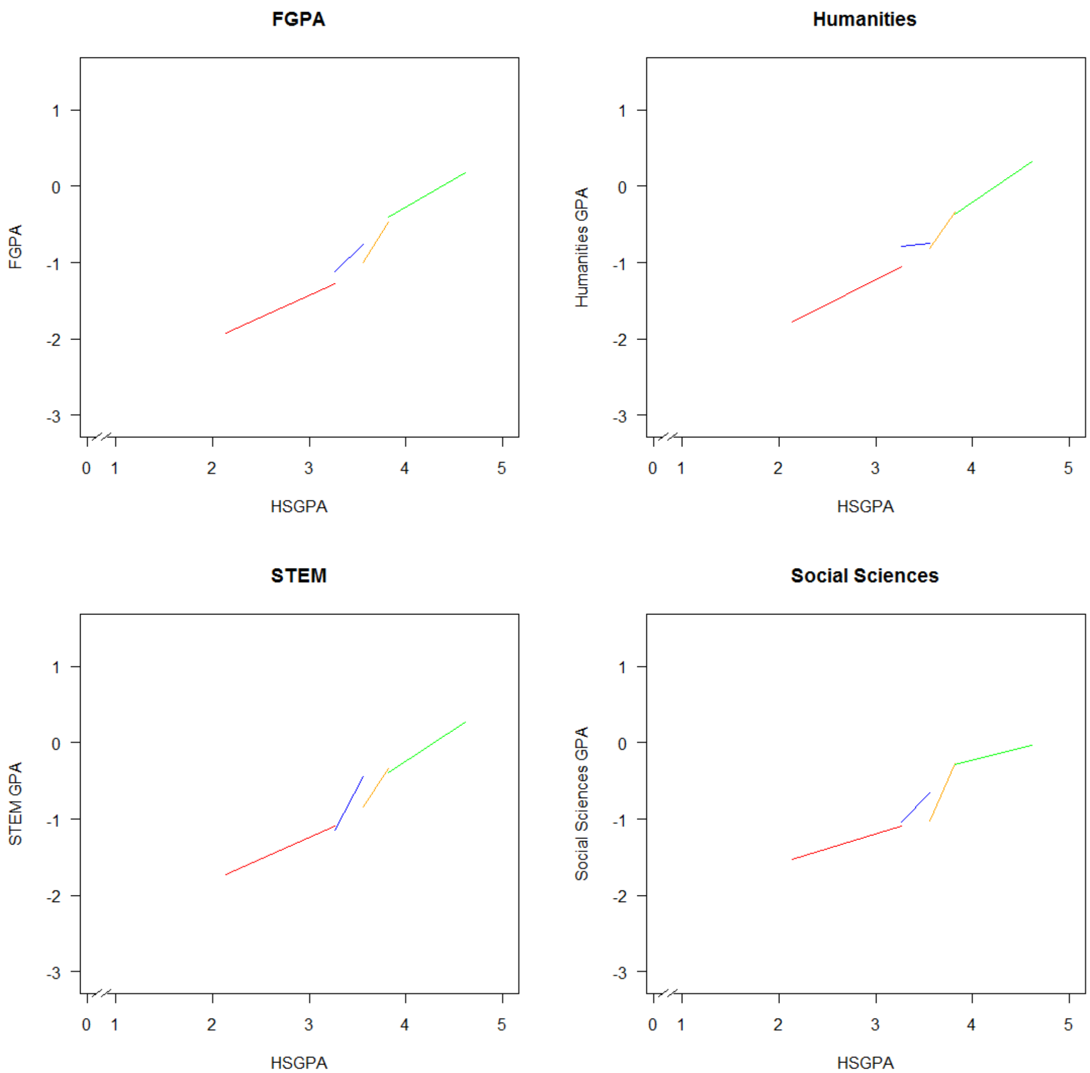

Humanities
Humanities GPA
HSGPA
STEM
STEM GPA
HSGPA
Social Sciences
Social Sciences GPA
HSGPA

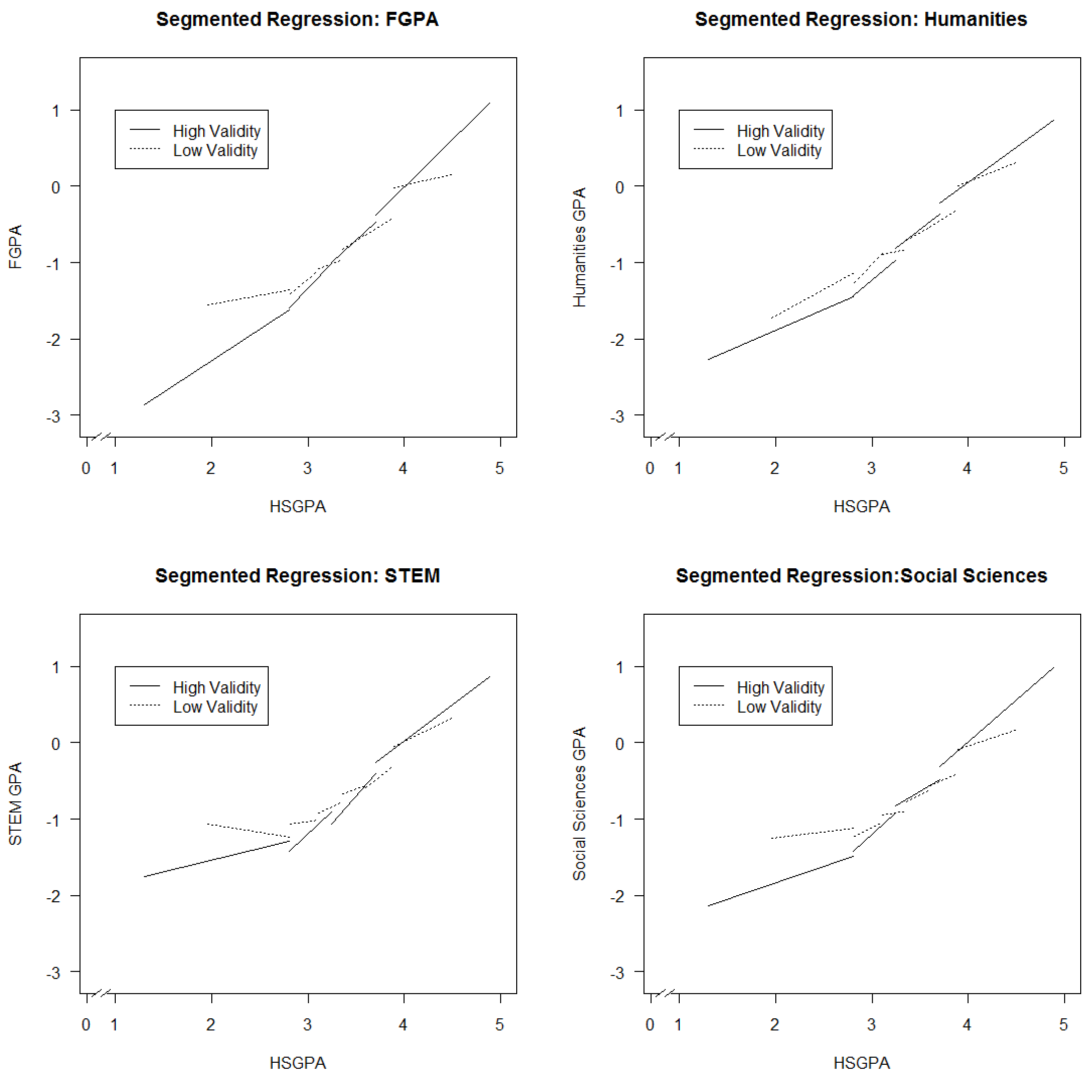

FGPA
Humanities
STEM
Social Sciences
FGPA
Humanities GPA
STEM GPA
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5

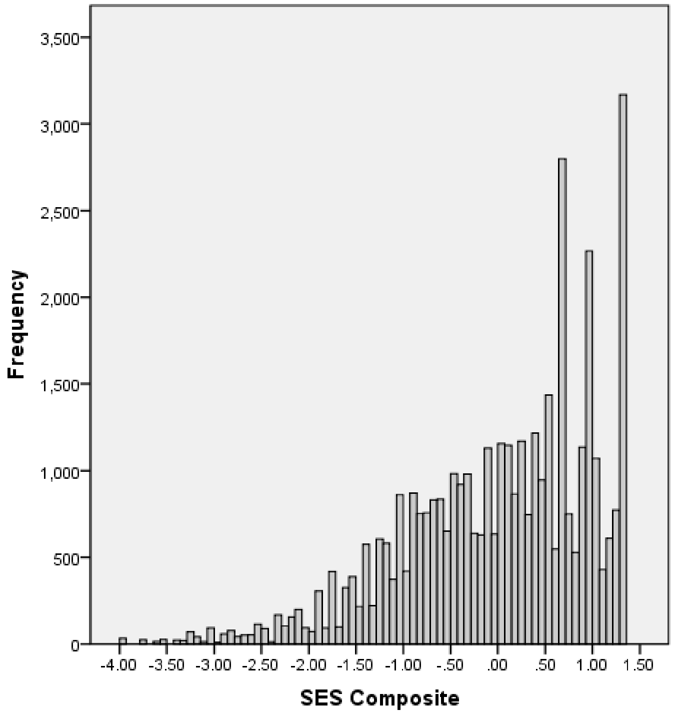

FGPA
FGPA
HSGPA
Humanities
Humanities GPA
HSGPA
STEM
STEM GPA
HSGPA
Social Sciences
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

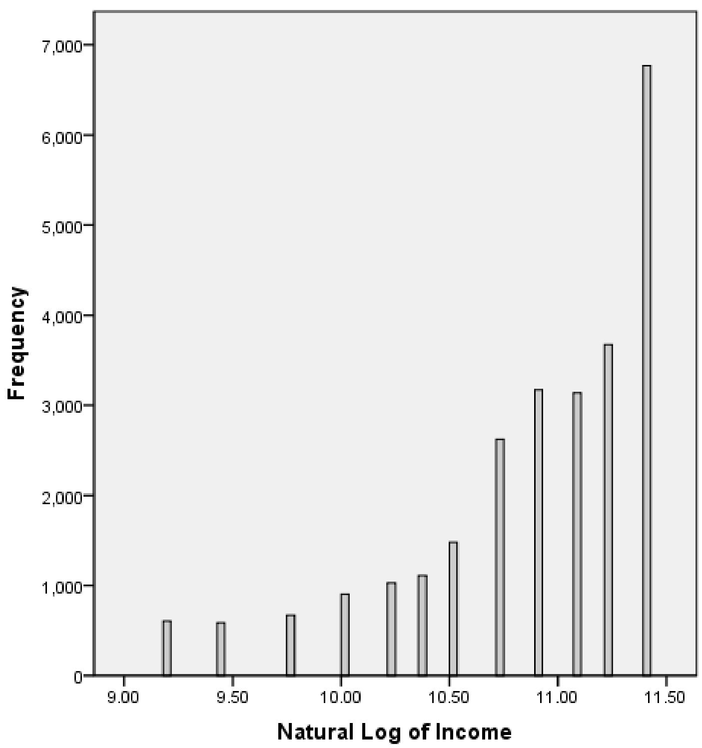

FGPA
Humanities
STEM
Social Sciences
FGPA
Humanities GPA
STEM GPA
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

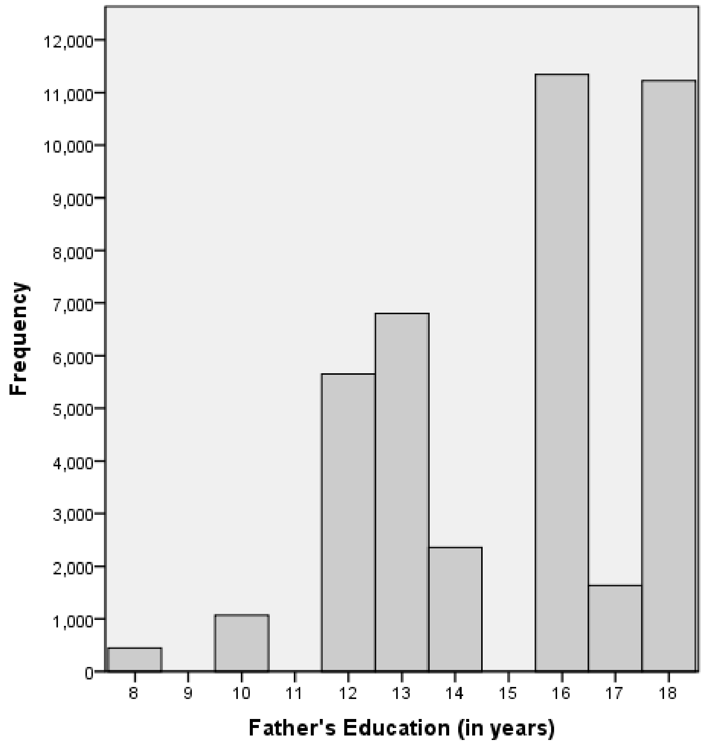

FGPA
Humanities
STEM
Social Sciences
FGPA
Humanities GPA
STEM GPA
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

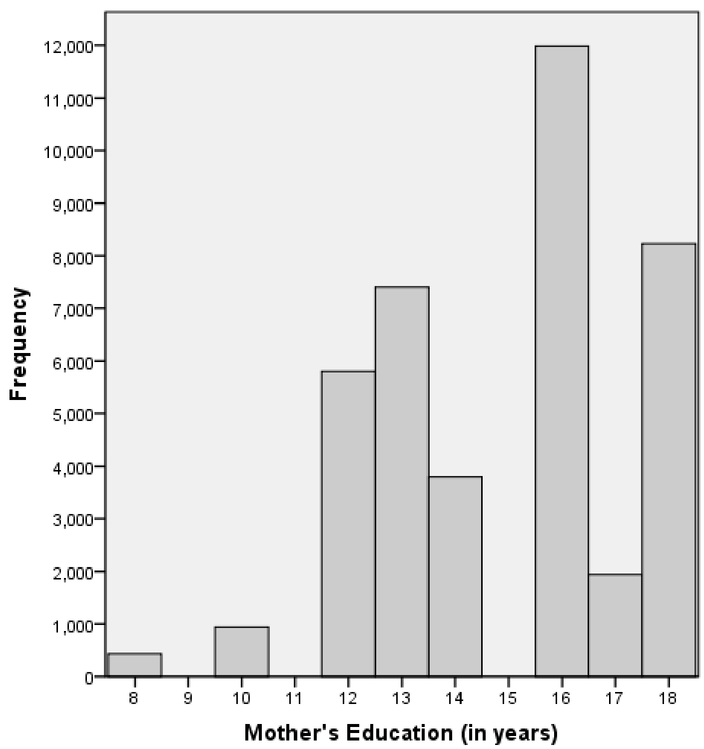

FGPA
FGPA
HSGPA
Humanities
Humanities GPA
HSGPA
STEM
STEM GPA
HSGPA
Social Sciences
Social Sciences GPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

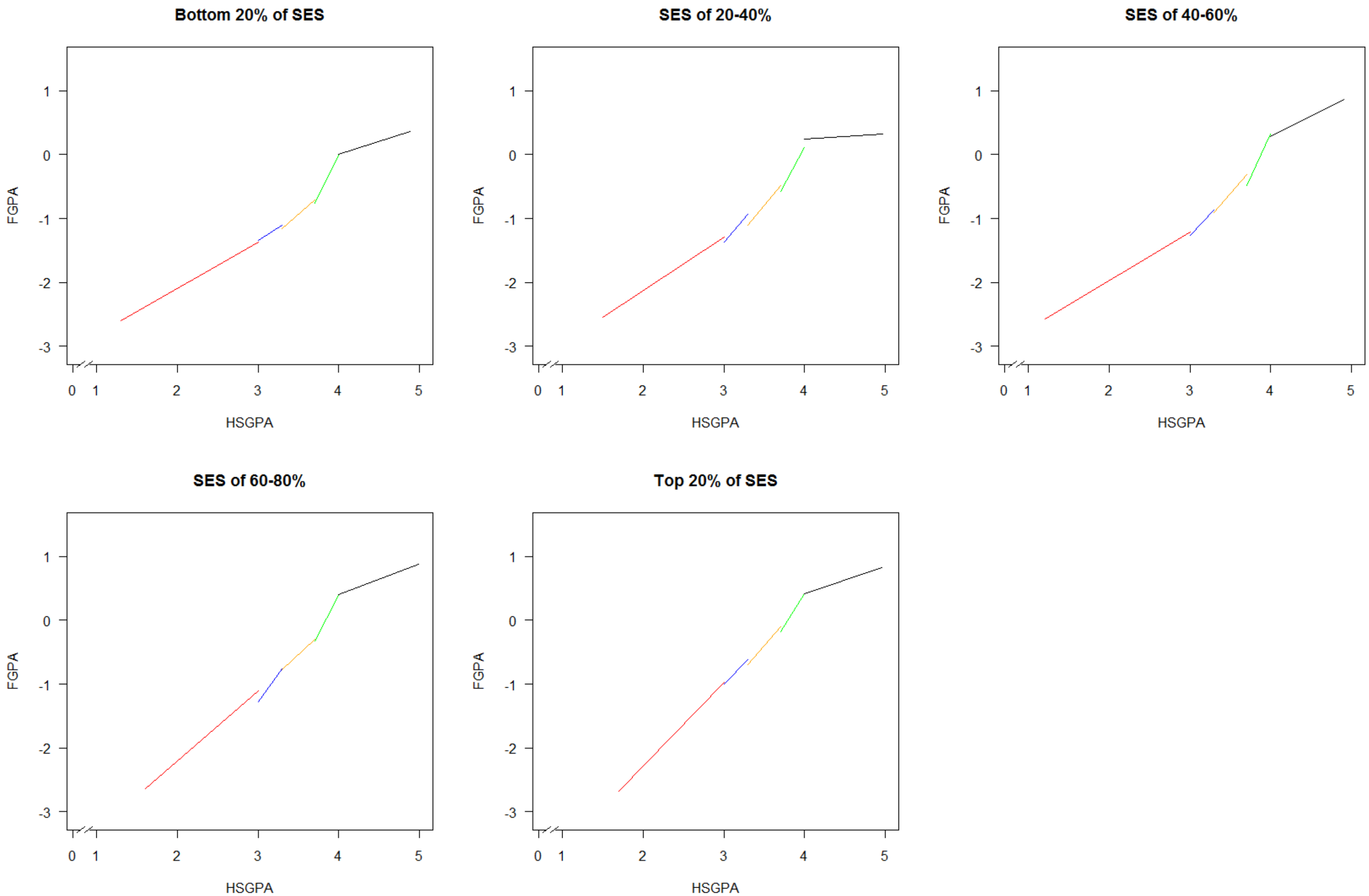

Segmented Regression: FGPA
FGPA
HSGPA
High Validity
Low Validity
Segmented Regression: Humanities
Humanities GPA
Segmented Regression: STEM
STEM GPA
Segmented Regression:Social Sciences
Social Sciences GPA

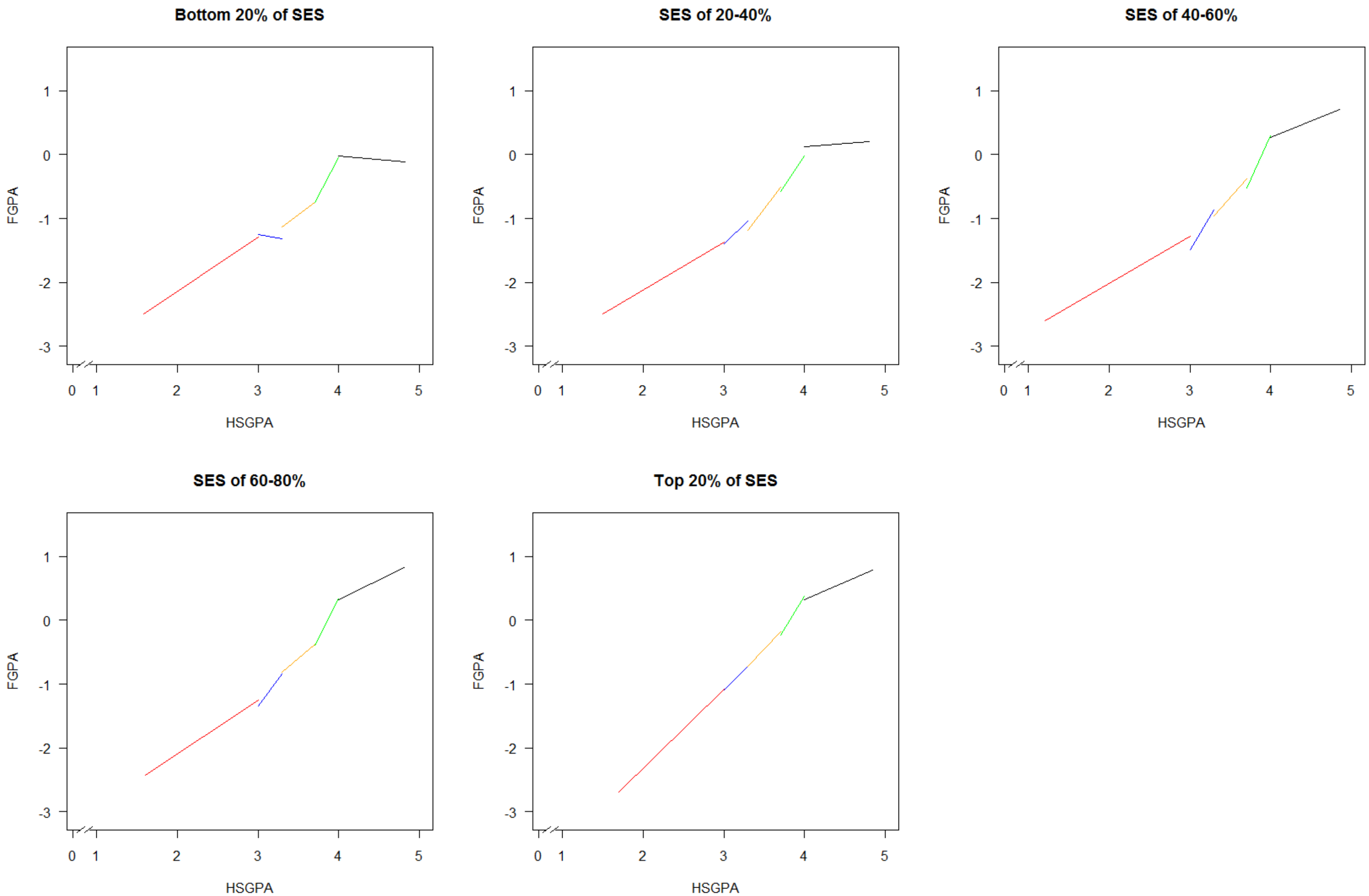

Frequency
3,500
3,000
2,500
2,000
1,500
1,000
500
0
-4.00 -3.50 -3.00 -2.50 -2.00 -1.50 -1.00 -.50 .00 .50 1.00 1.50
SES Composite


Frequency
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
9.00 9.50 10.00 10.50 11.00 11.50
Natural Log of Income

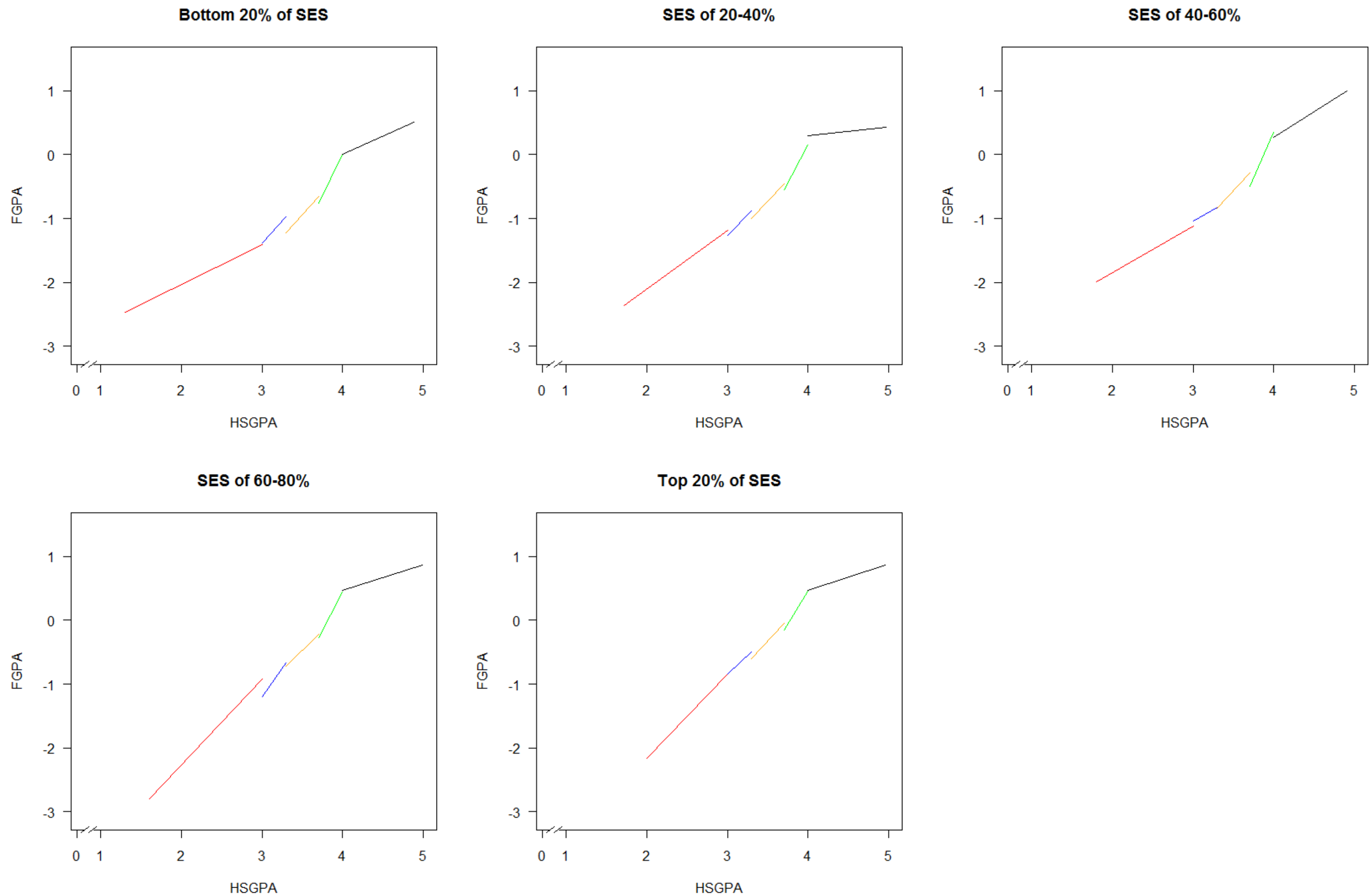

Frequency
12,000
11,000
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
8 9 10 11 12 13 14 15 16 17 18
Father's Education (in years)


Frequency
12,000
11,000
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
8 9 10 11 12 13 14 15 16 17 18
Mother's Education (in years)

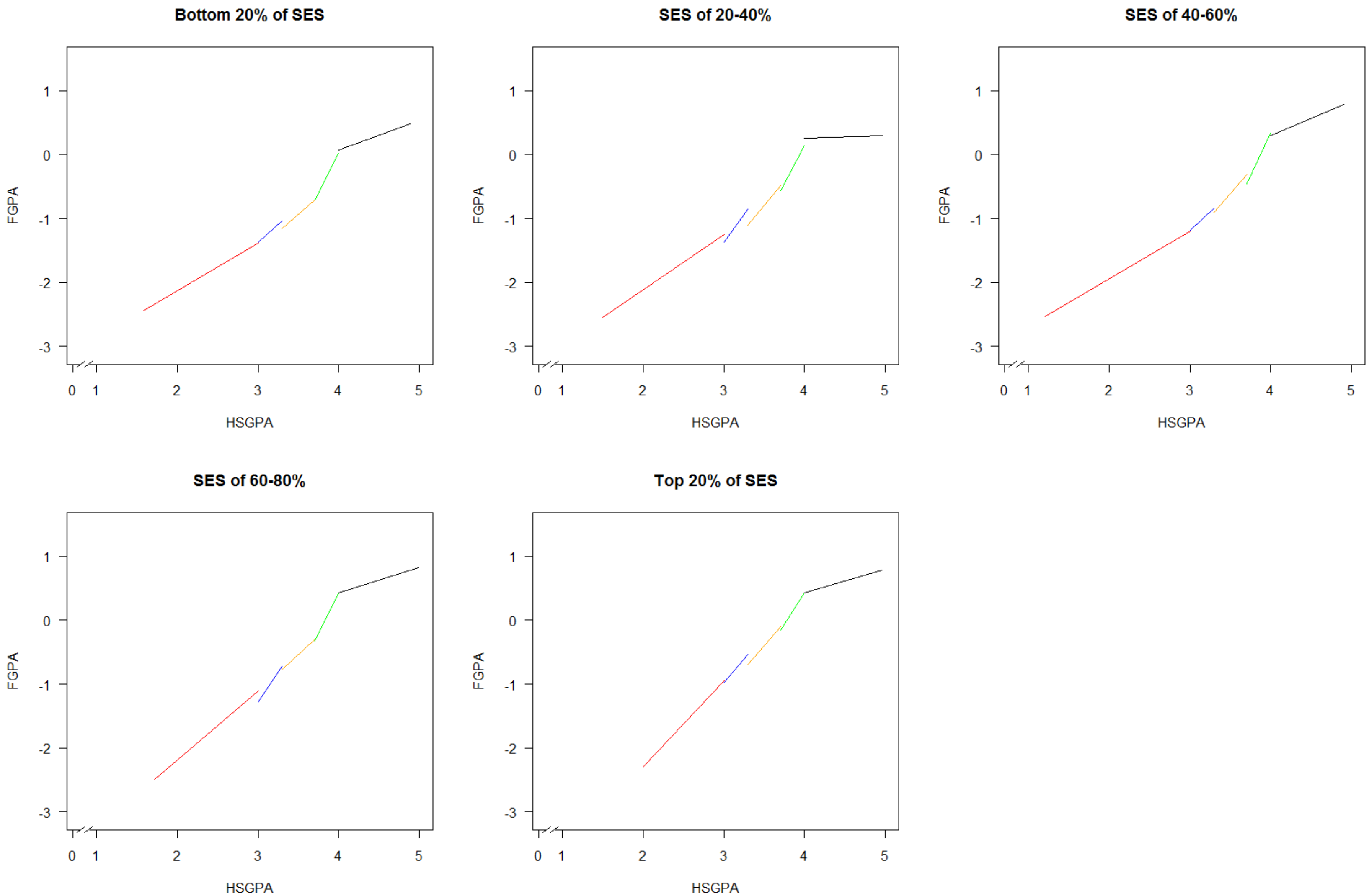

Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

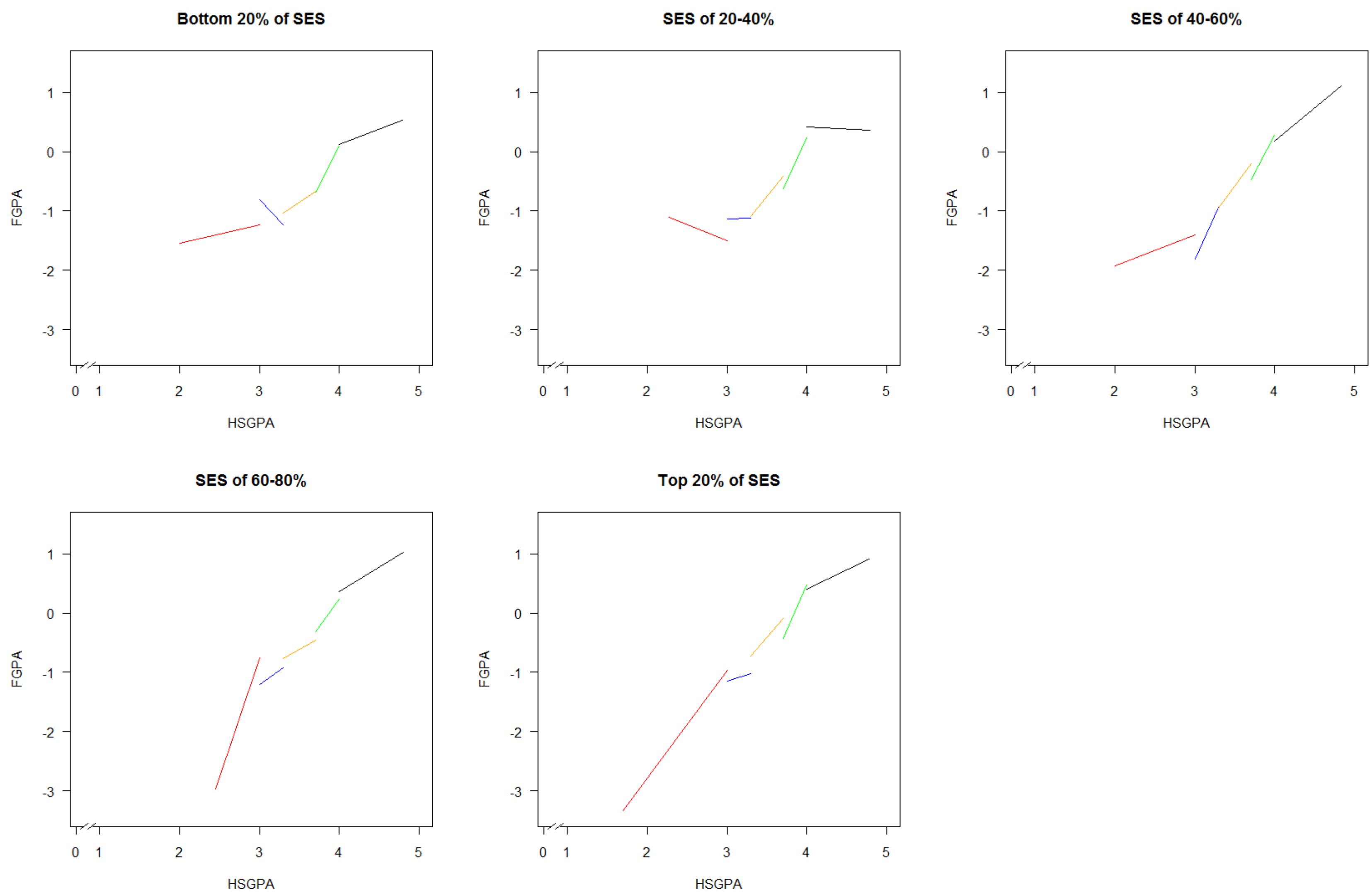


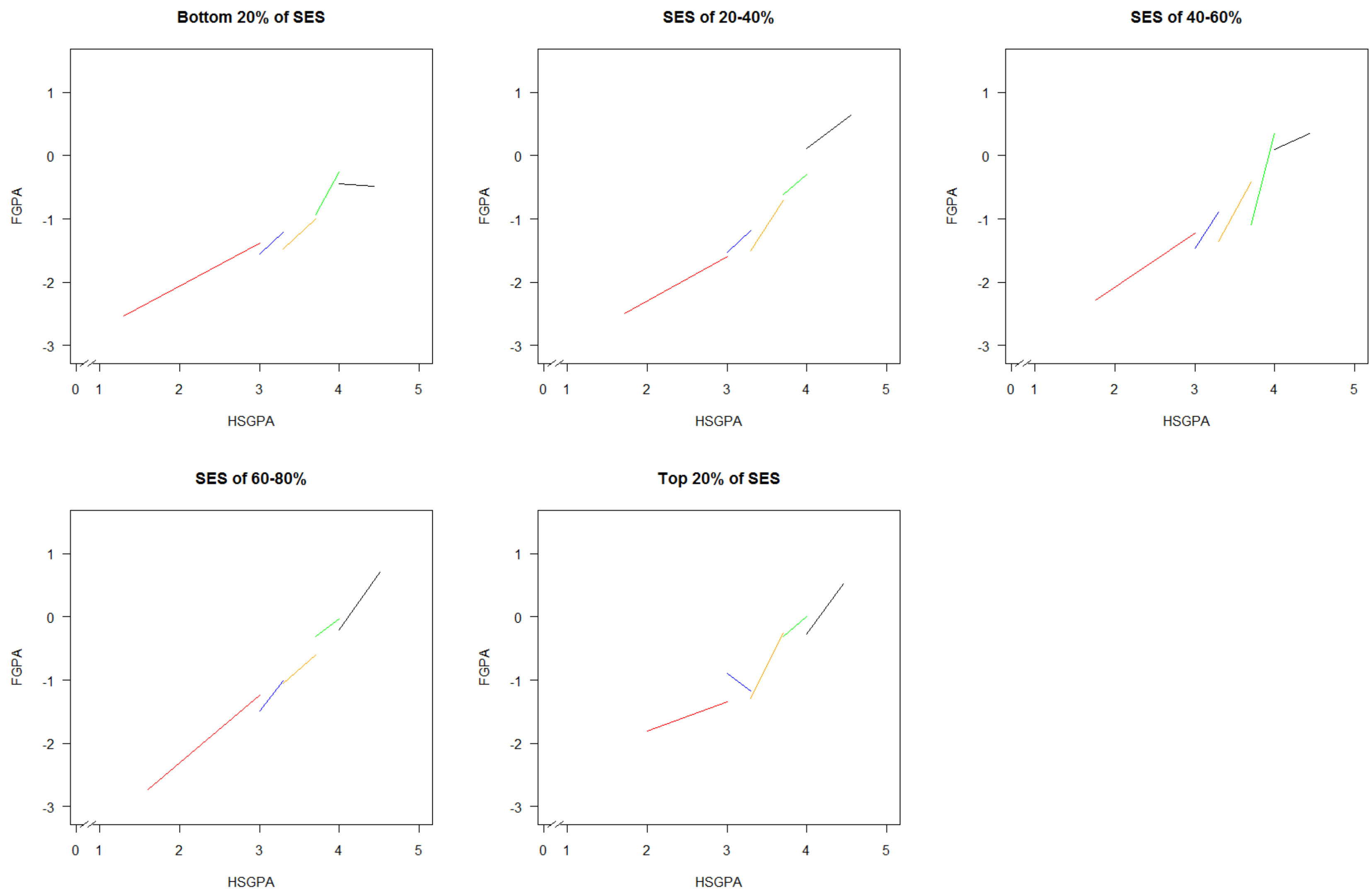
Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA

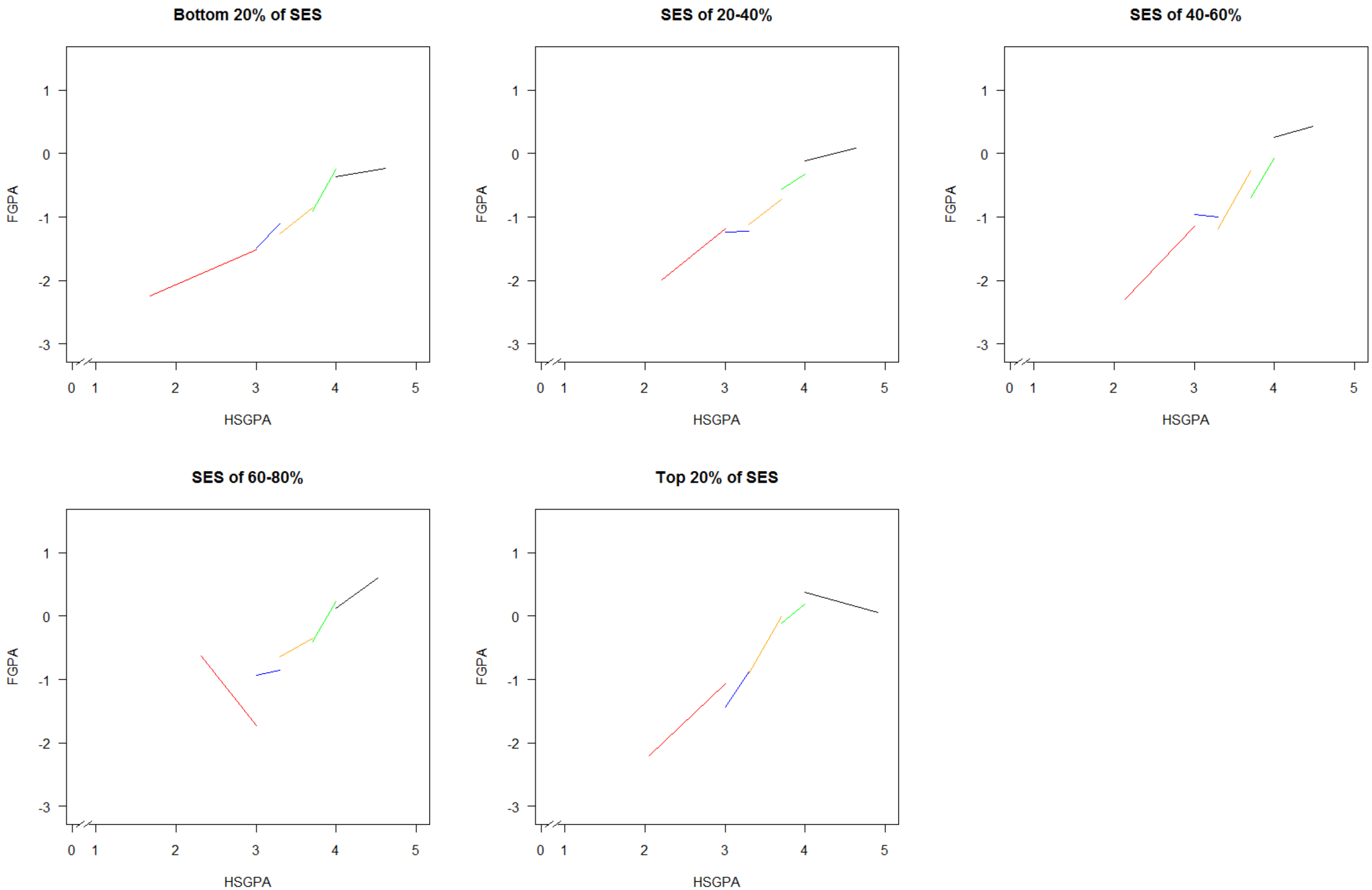


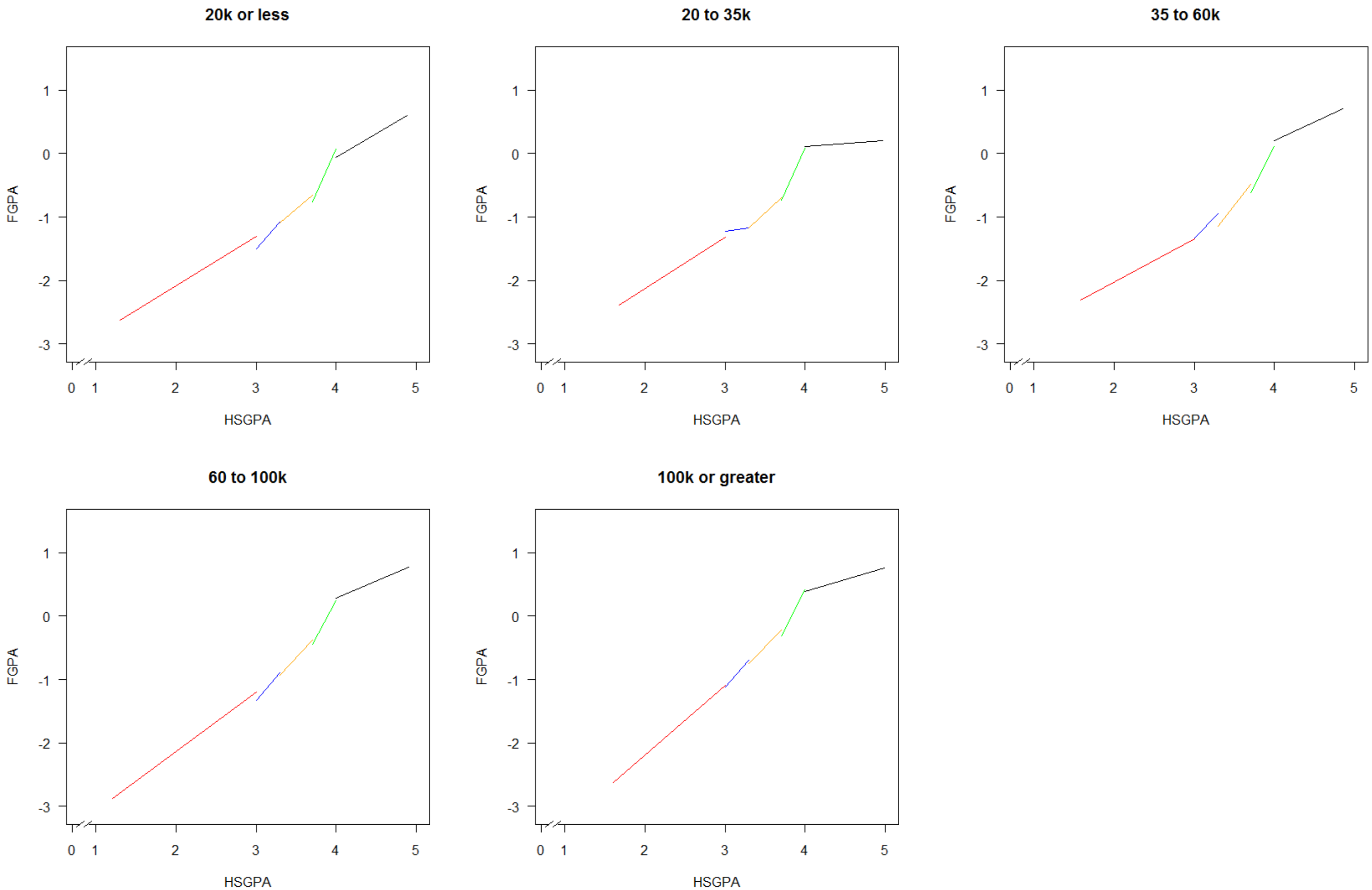
Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA

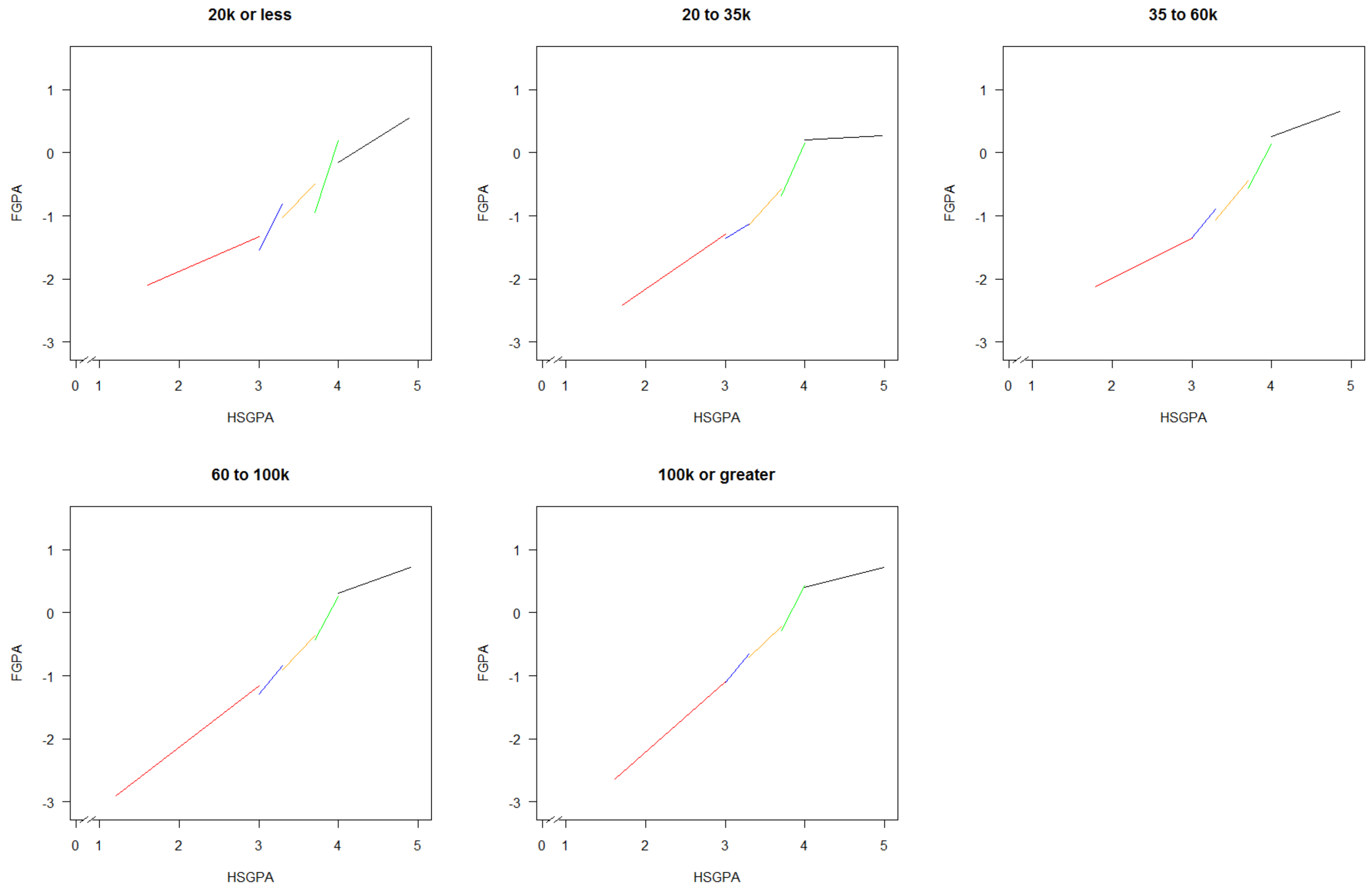

Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
1
0
-1
-2
-3
0
1
2
3
4
5
HSGPA

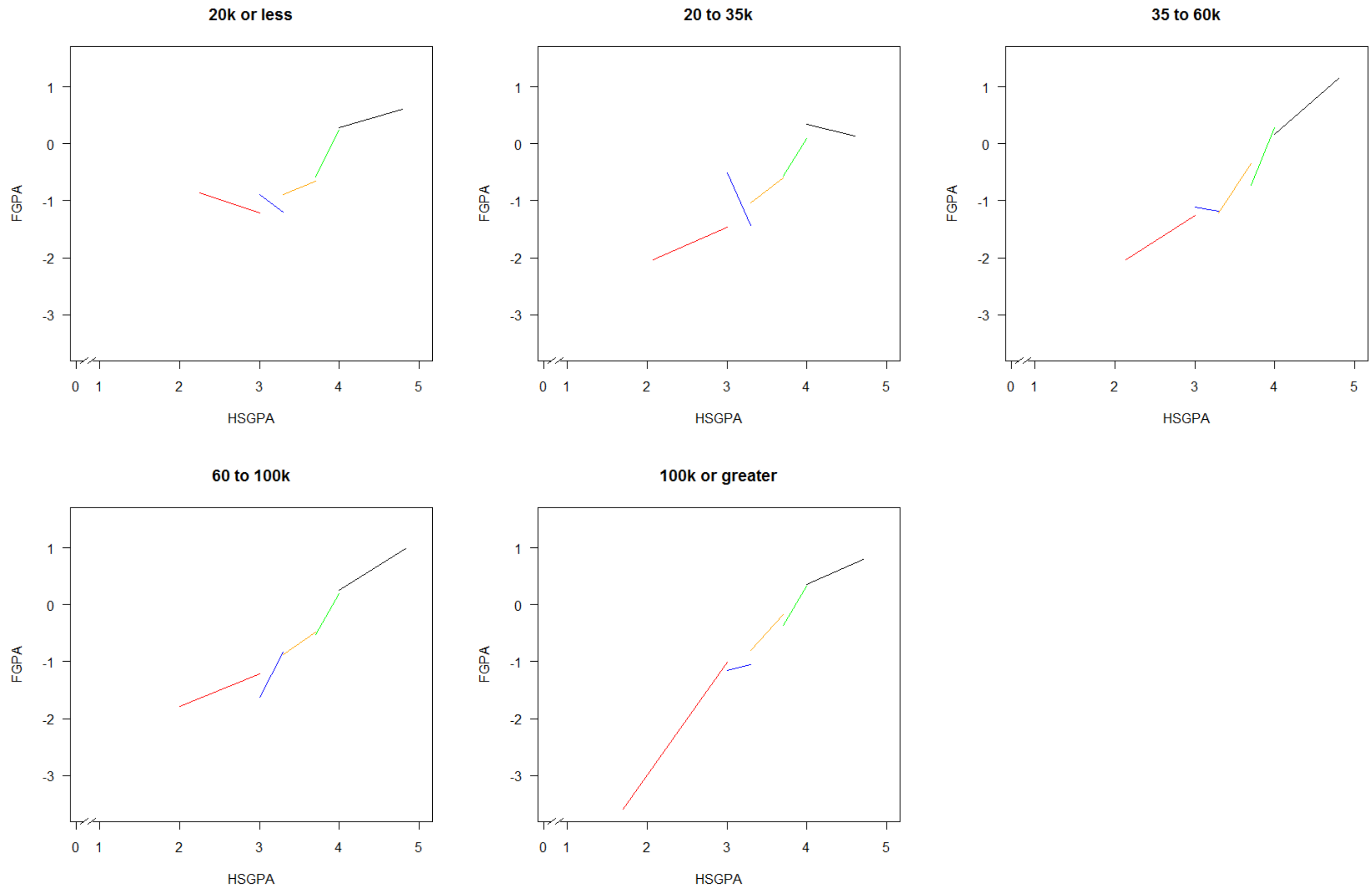

Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA

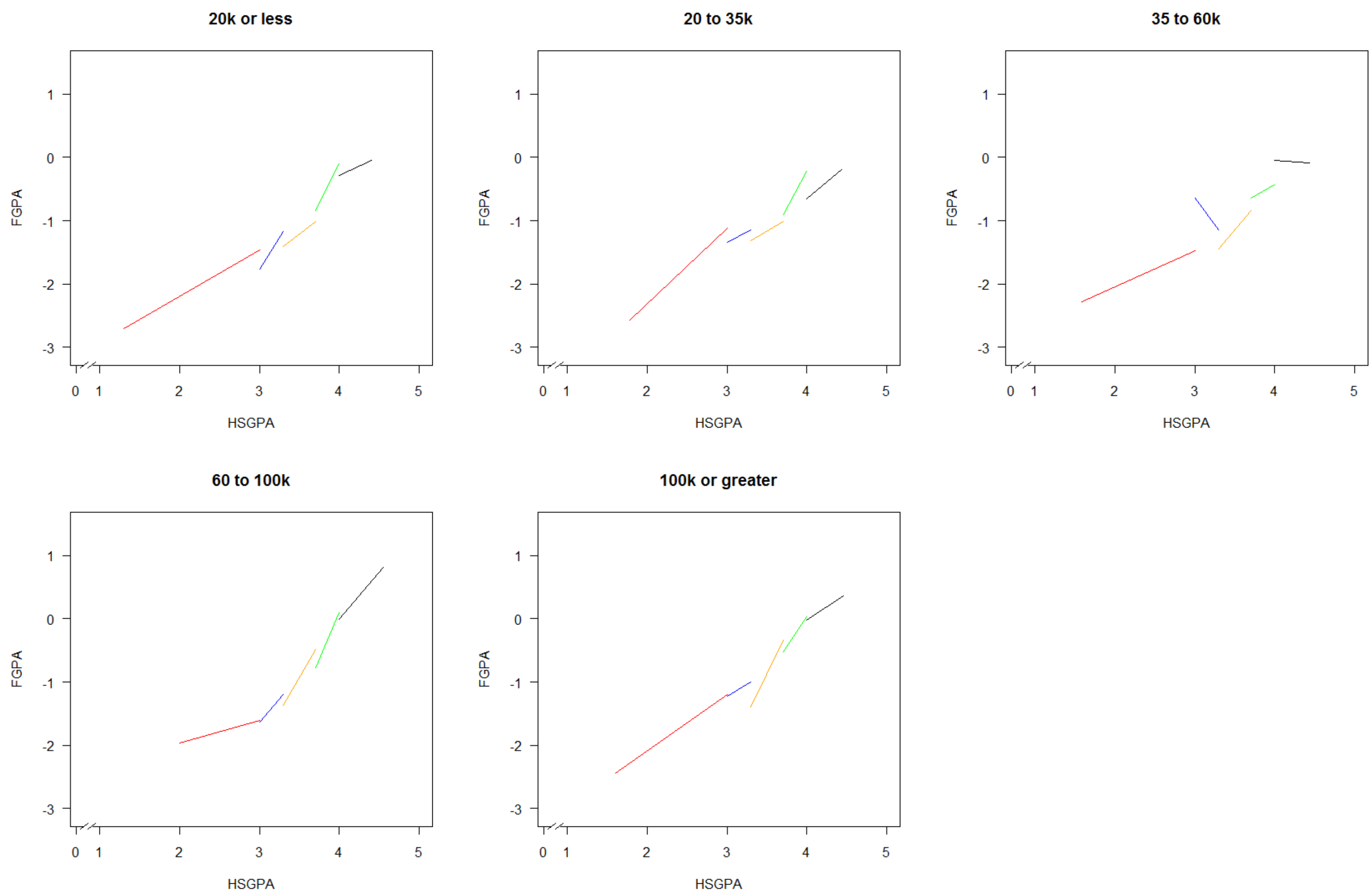

Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA


Bottom 20% of SES
SES of 20-40%
SES of 40-60%
SES of 60-80%
Top 20% of SES
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

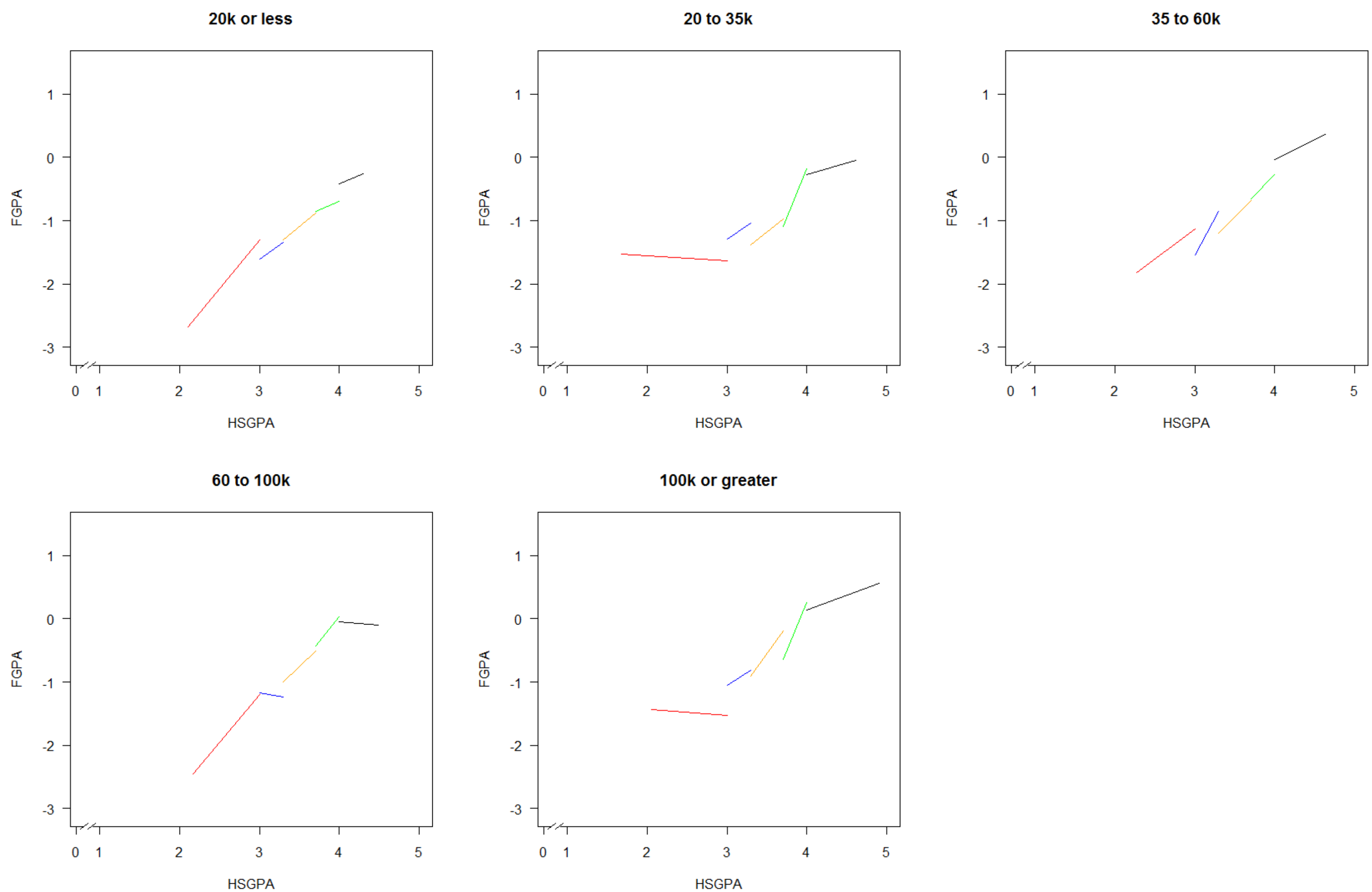

20k or less
20 to 35k
35 to 60k
60 to 100k
100k or greater
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

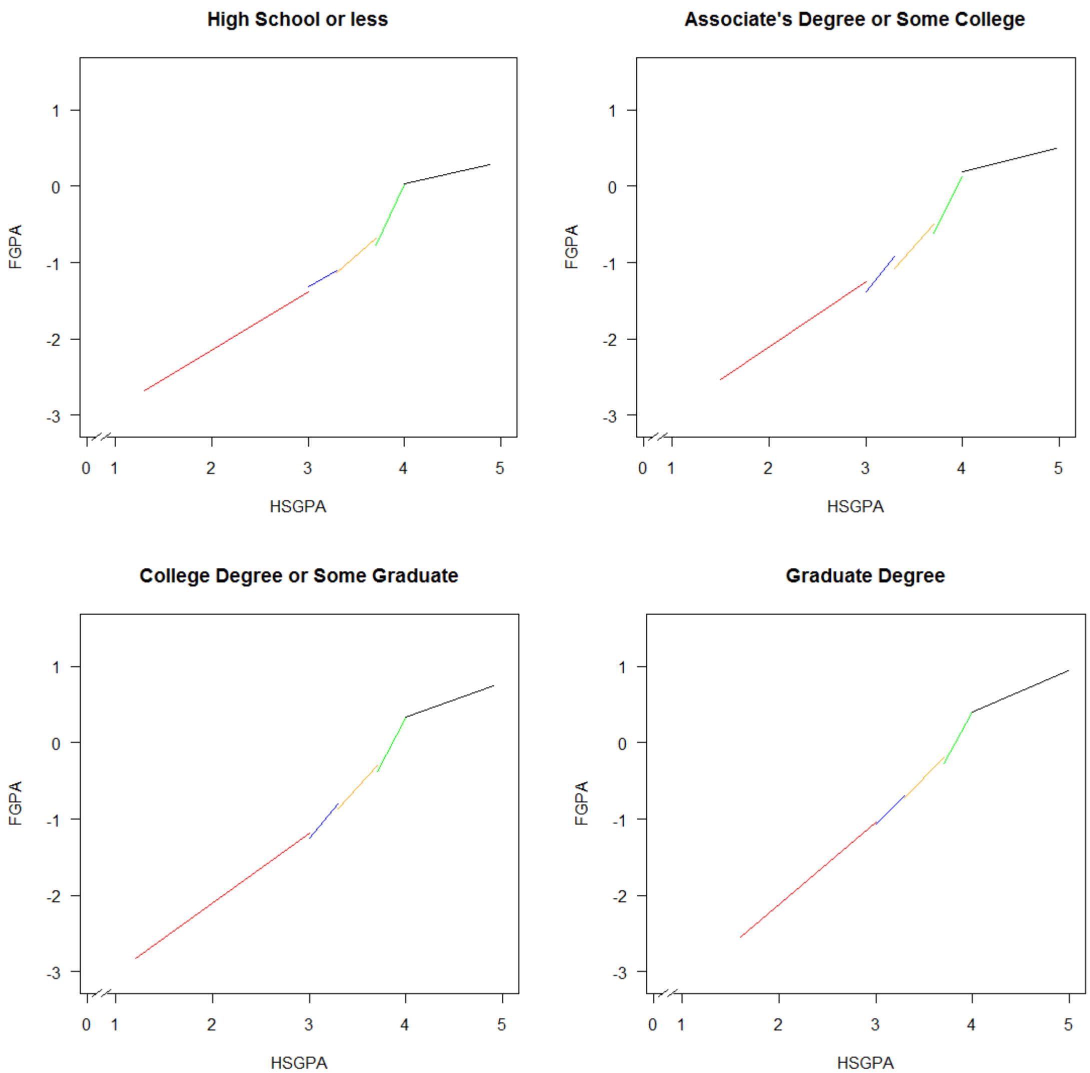

20k or less
20 to 35k
35 to 60k
60 to 100k
100k or greater
FGPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5

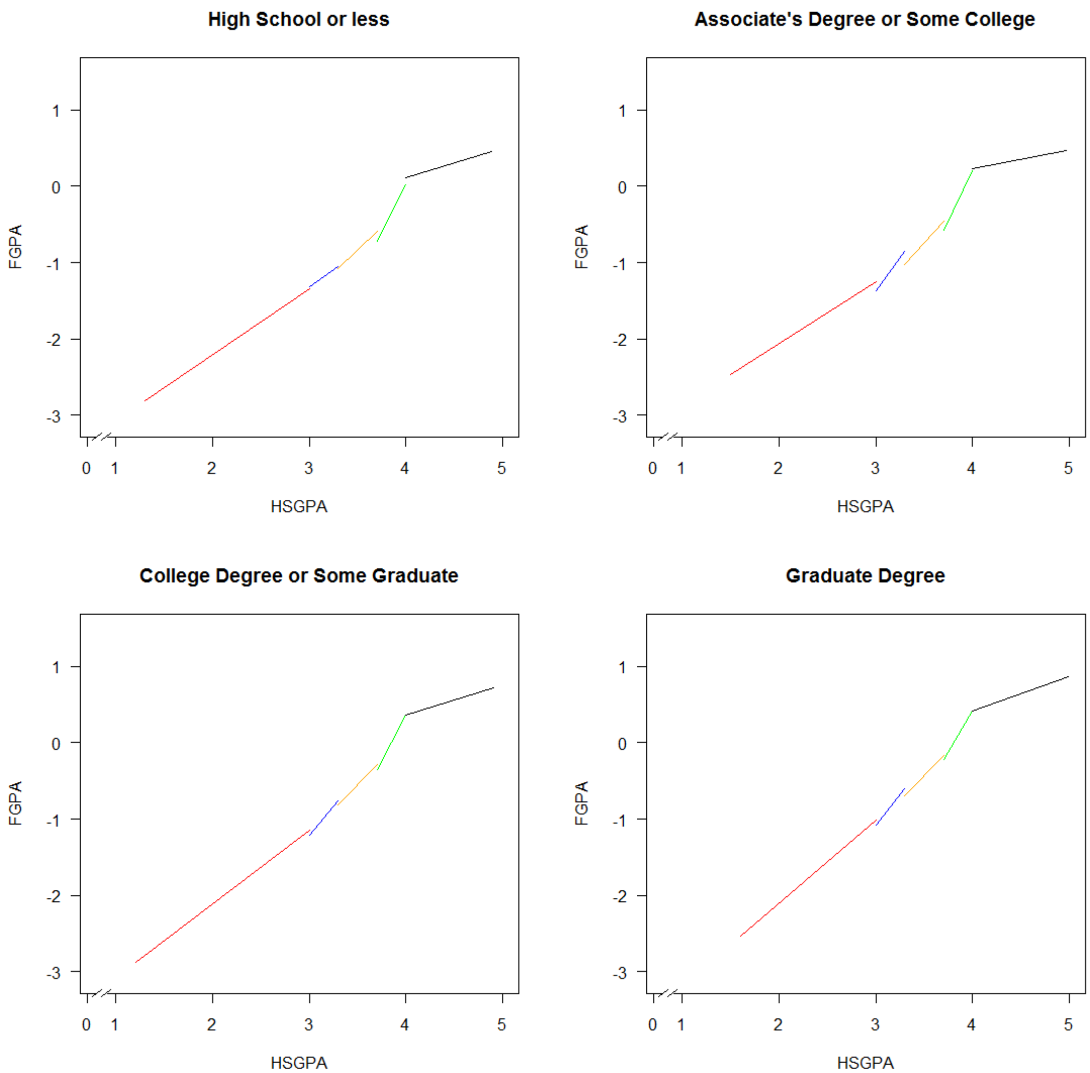

20k or less
20 to 35k
35 to 60k
60 to 100k
100k or greater
FGPA
HSGPA

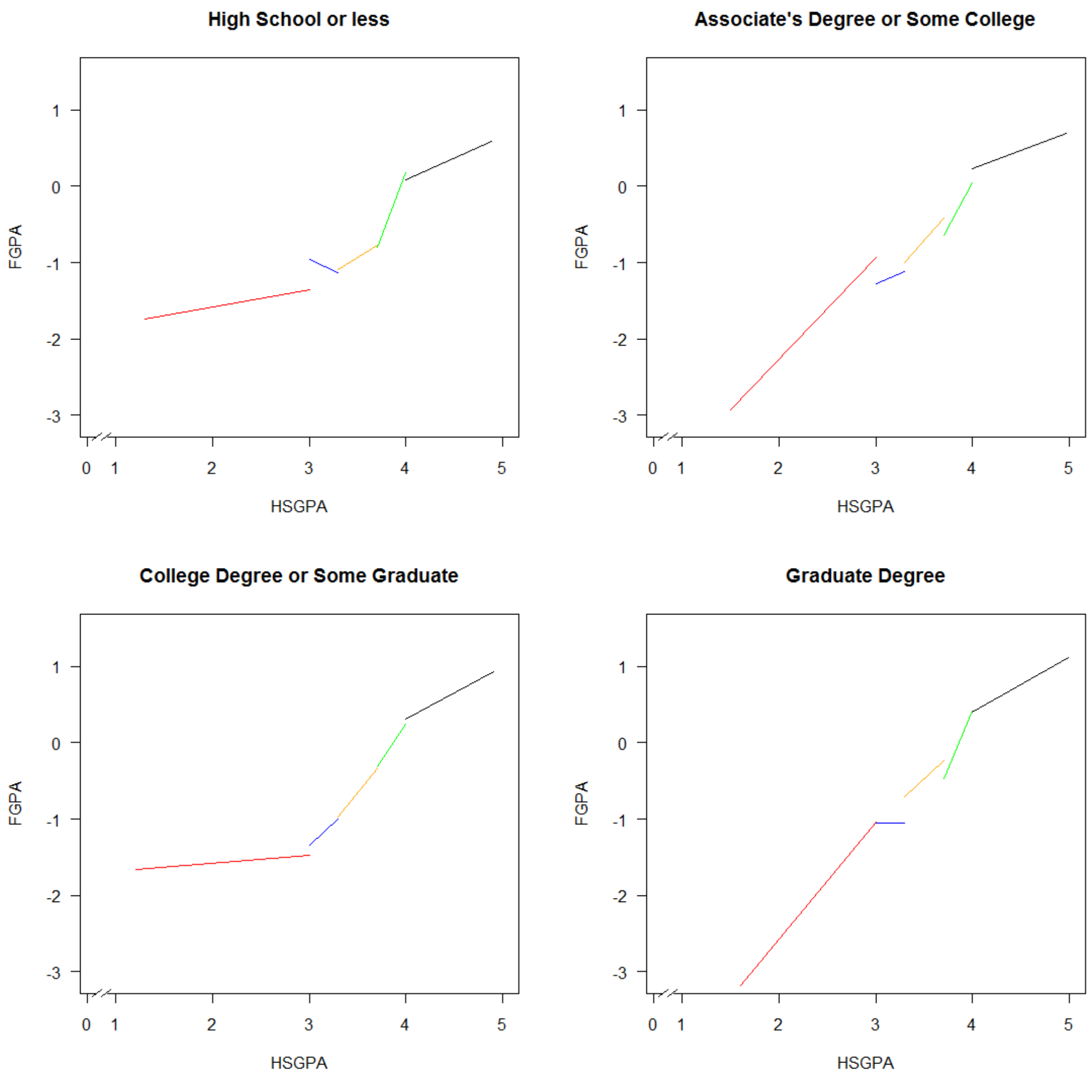


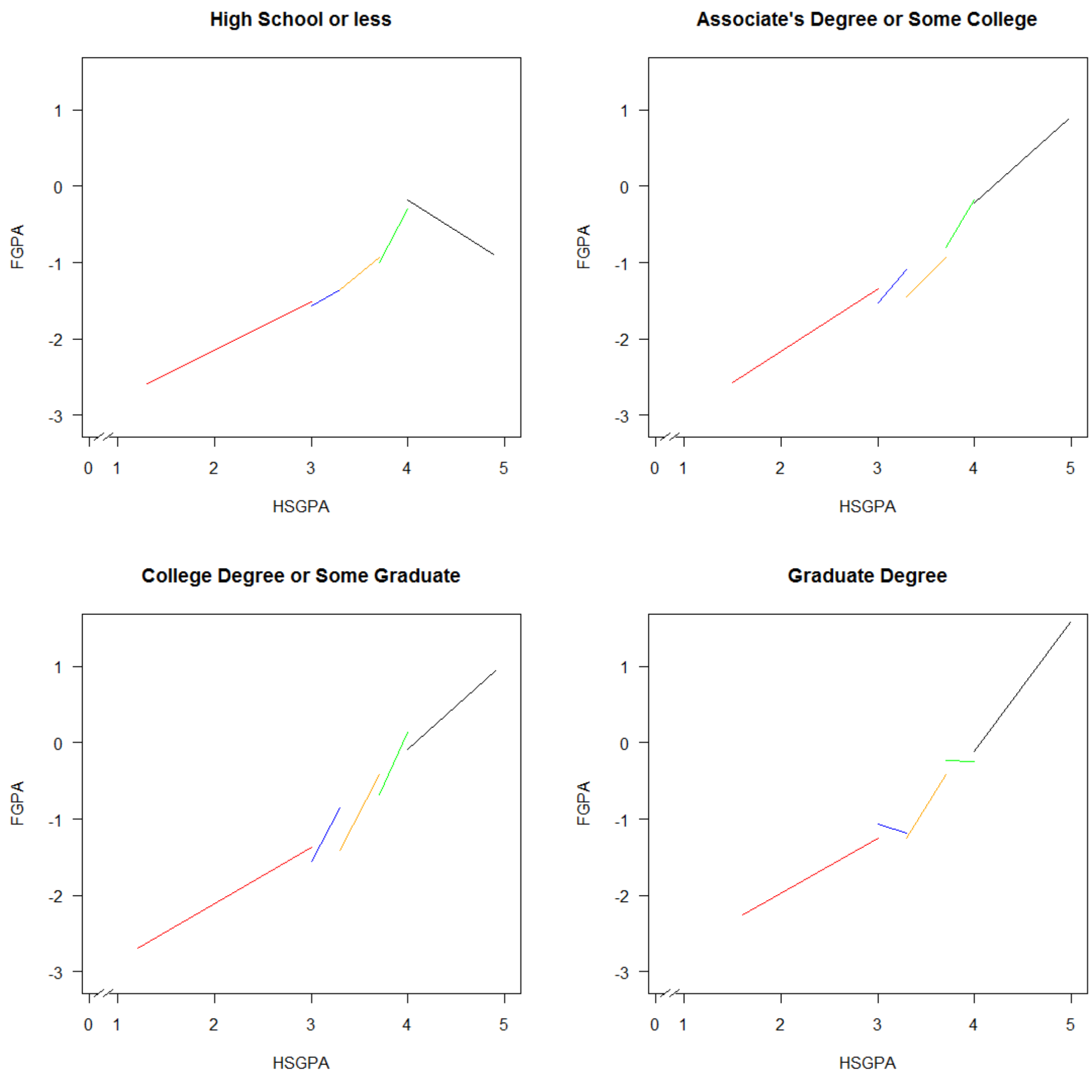
20k or less
20 to 35k
35 to 60k
60 to 100k
100k or greater
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

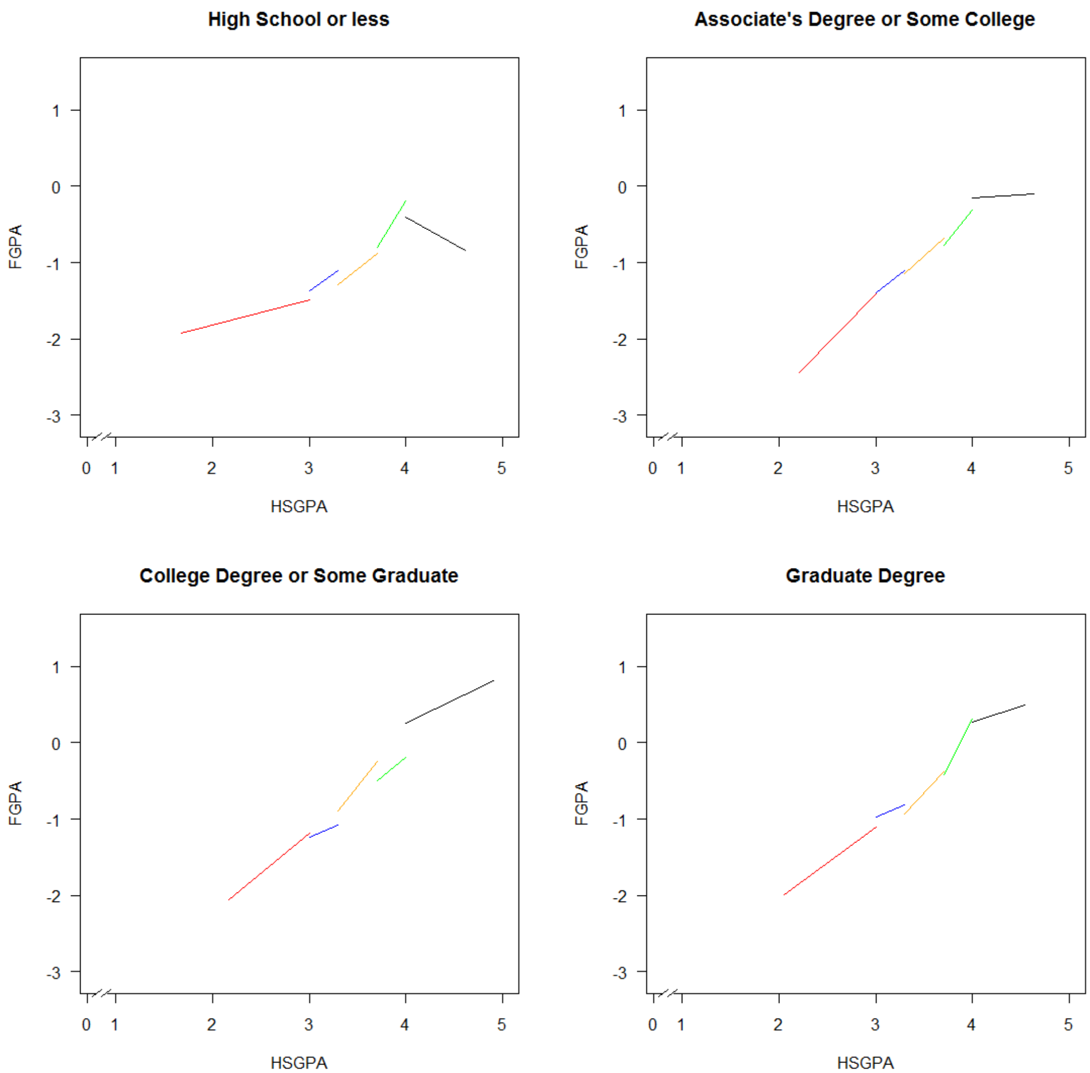

20k or less
20 to 35k
35 to 60k
60 to 100k
100k or greater
FGPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5

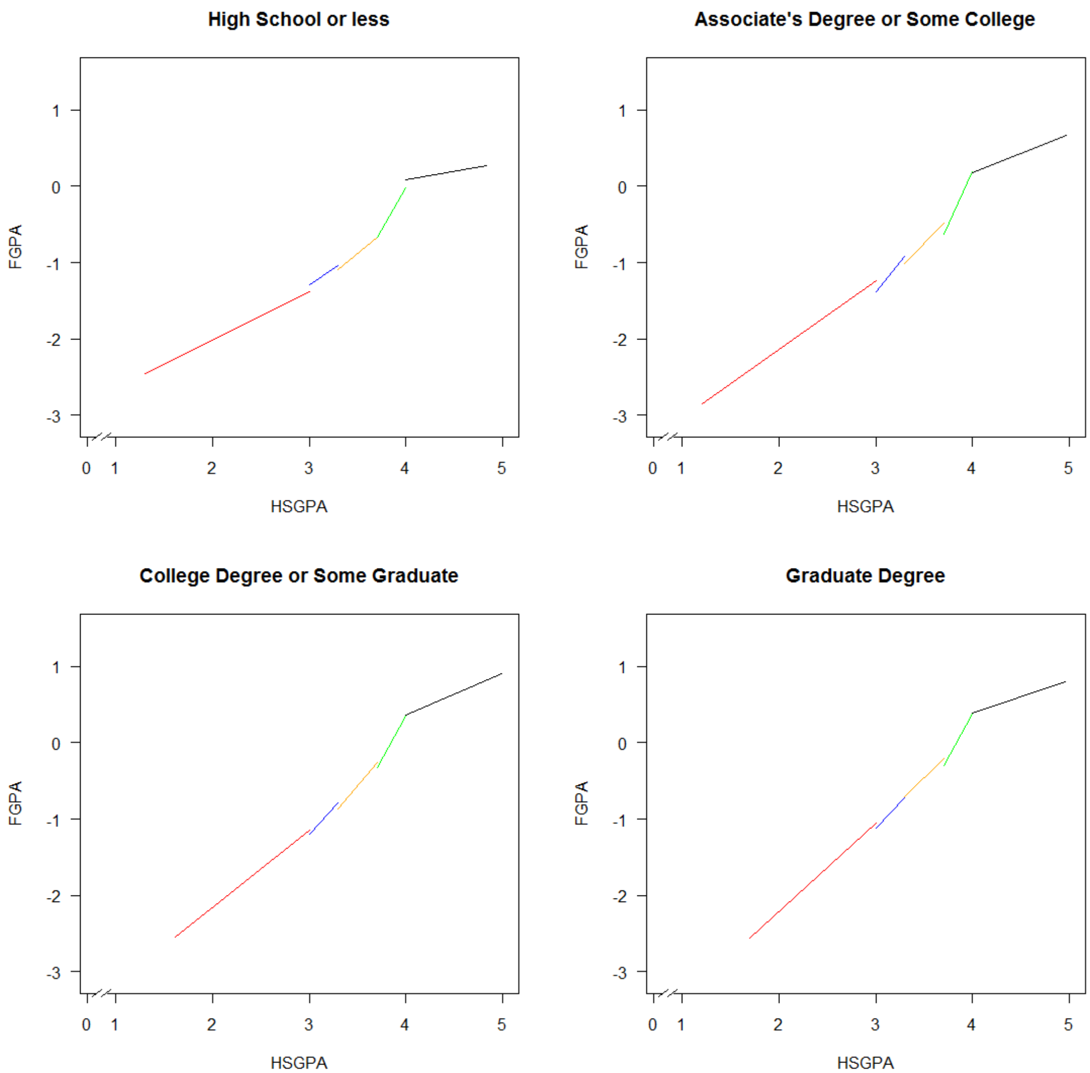

High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5

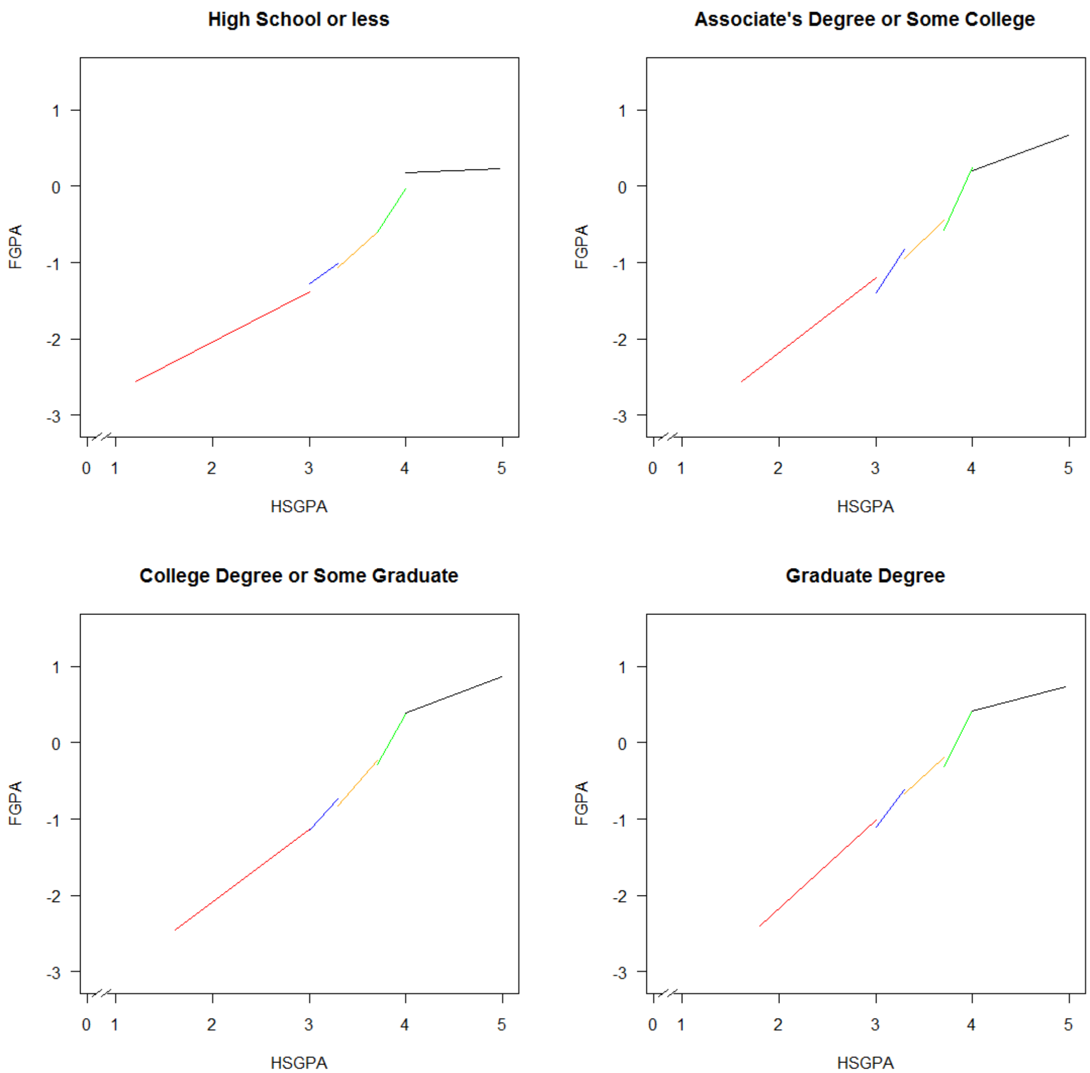

High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

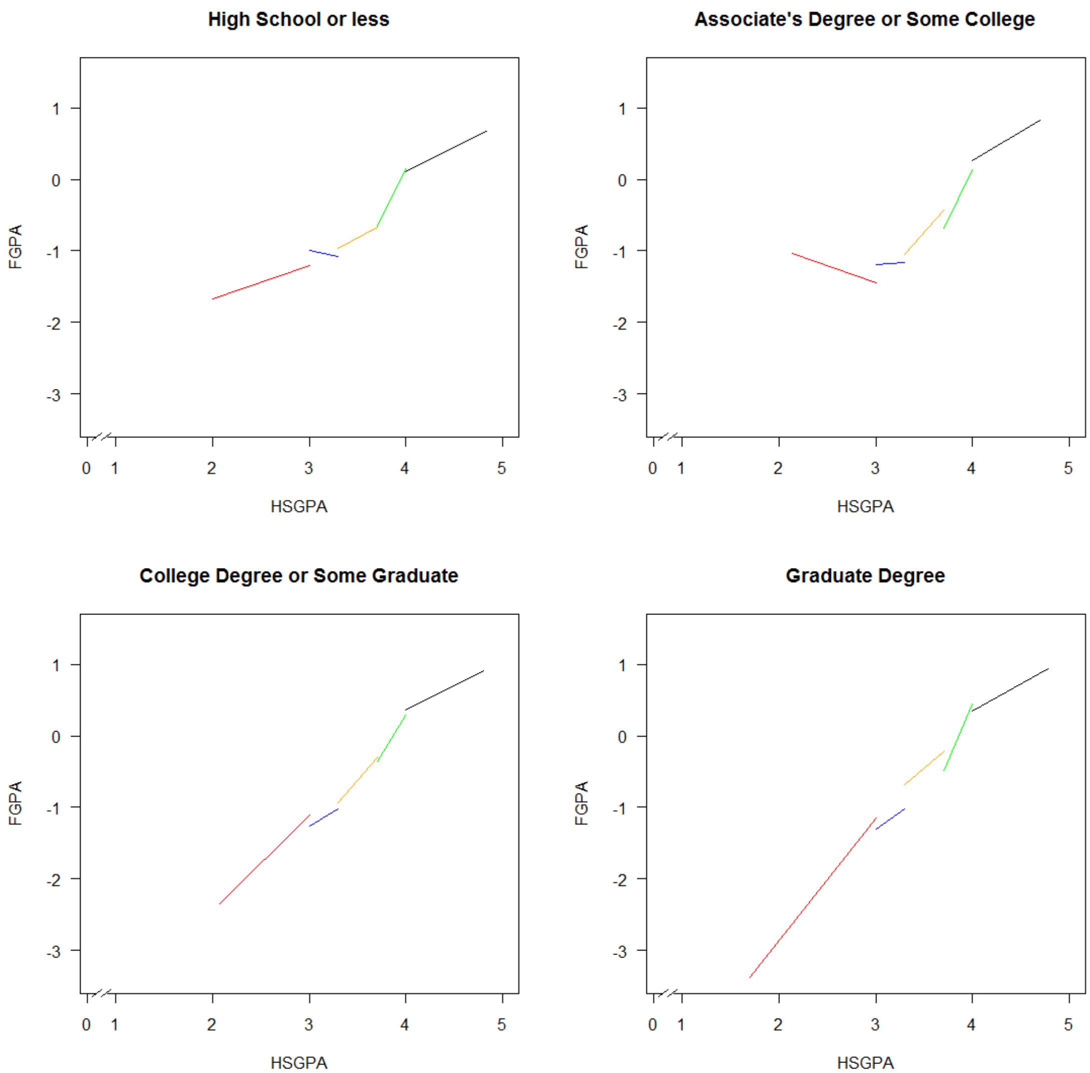

High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

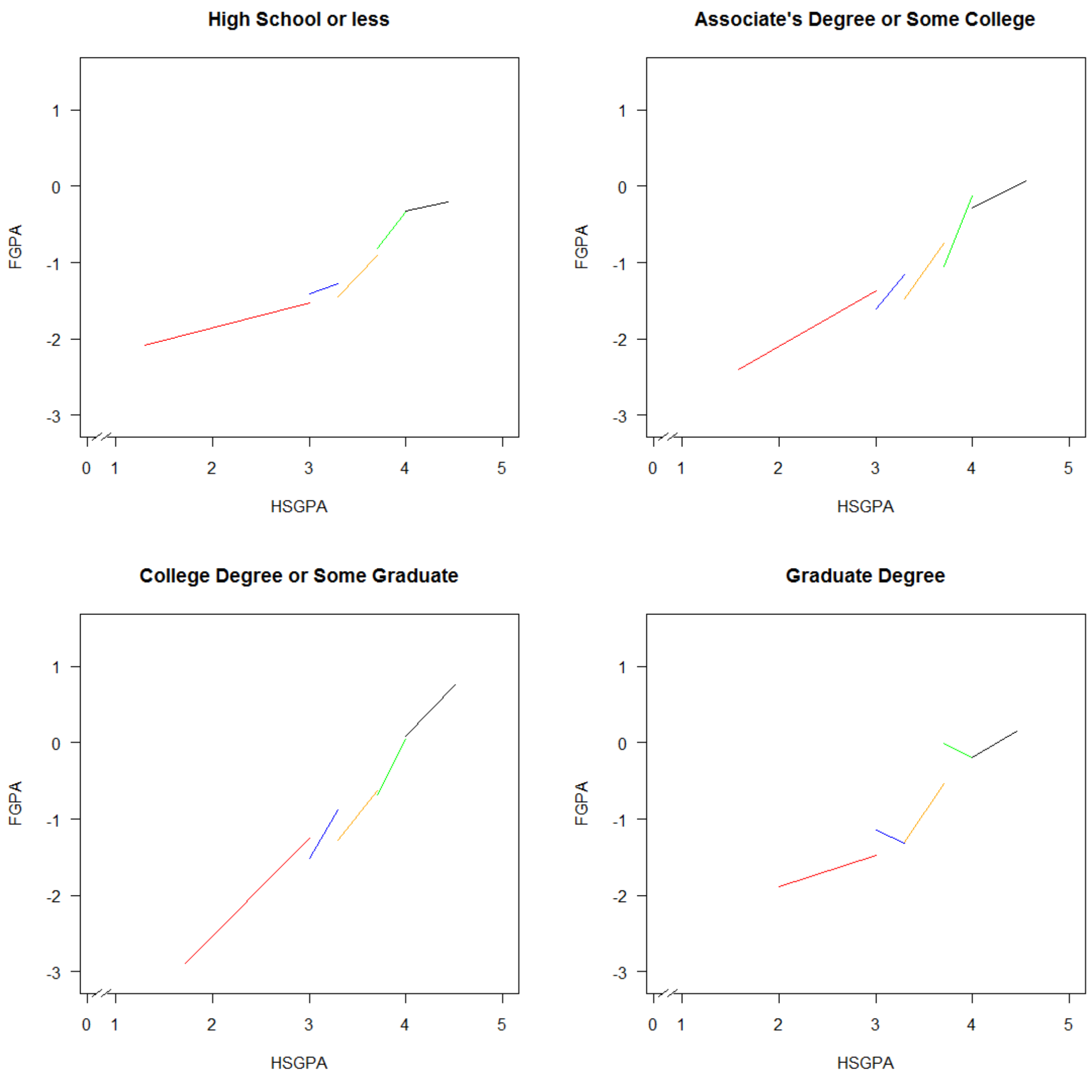

High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5

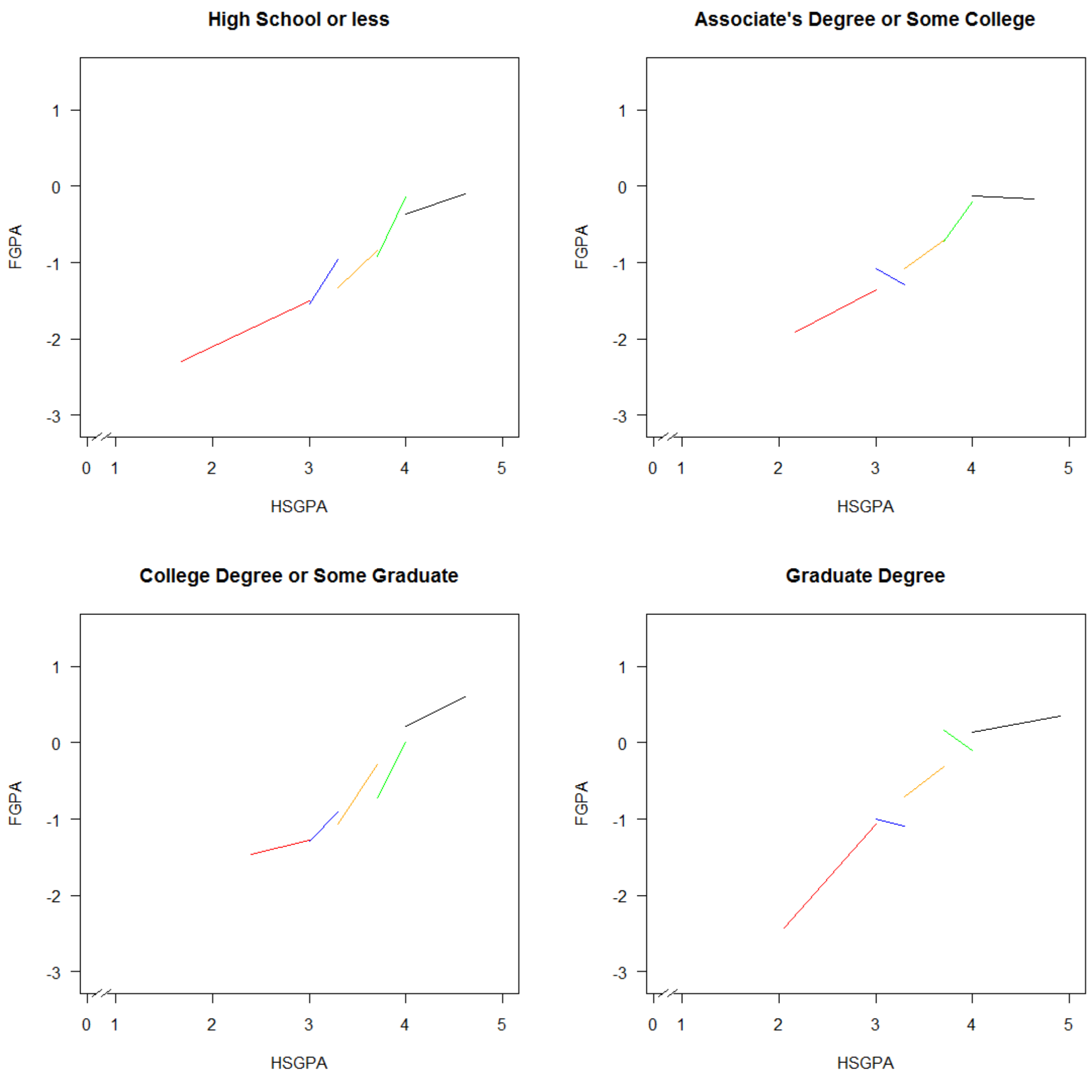

High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5


High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0 1 2 3 4 5


High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5


High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5


High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5


High School or less
Associate's Degree or Some College
College Degree or Some Graduate
Graduate Degree
FGPA
HSGPA
1
0
-1
-2
-3
0
1
2
3
4
5