Discussion Paper No. 69

Williams Project on the Economics of Higher Education
23 Whitman Street
Mears West
Williams College
Williamstown, MA 01267

http://www.williams.edu/wpehe

Access to the Most Selective Private[1] Colleges by High-Ability, Low-Income Students: Are they out there?

Gordon C. Winston
Catharine B. Hill[2]
Williams College

October, 2005
DP-69

[1] Public institutions are omitted for want of data.

[2] MiHye Kim got this project underway with her examination of the SAT data and the momentum was continued by Rachel Louis, Ashley Hartman, Carey Aubert, and Jessica England. David Brodigan gave us access to Williams' data. The cooperation of the College Board and ACT in providing data was crucial and is much appreciated. Very helpful comments came from Dave Zimmerman, Robert Klitgaard, Chris Winters and participants at the Macalester-Spencer Forum in June 2005 and the Forum for the Future of Higher Education at Aspen in September. As usual, the Andrew W. Mellon Foundation's support of the Williams Project on the Economics of Higher Education was critical in making this study possible.

ABSTRACT

With only a small number of their students coming from families with the lowest incomes (10% from the bottom two family income quintiles), the nation's most selective private colleges and universities need to know why. Two ready ideological answers are (1) that low-income high-ability students are being excluded in order to favor the children of society's most advantaged or (2) that very few low-income high-ability students exist – that by college age, low-income students have been so damaged by education, nutrition, neighborhoods, and families that few can qualify in a perfectly fair admissions process.

This paper uses the national population of high school test-takers in 2003 to examine the national distribution over family incomes of high-ability students (variously defined). With these data, two questions can be addressed. What would be the target share of low-income students at these schools if their student bodies were to mirror the national high-ability population? And, are they out there – do there exist enough such low-income, high-ability students to meet those targets?

It is shown that they are out there – that a somewhat larger share of the test-taking population is made up of high-ability, low-income students than are found in these schools and that their numbers make it feasible for the schools to increase their enrollments to target that national share. Because much depends on the definition of "high-ability" used, we consider alternative definitions but reach the same conclusion at any reasonable level (like a minimum combined SAT of 1300 or even 1420).

Introduction

The growing concern about access to highly selective colleges and universities[3] was heightened by a soon-to-be-published study of 28 of the most selective private schools in the US – "the COFHE schools[4]" – that showed that only 10% of their students come from the bottom 40% of the US family income distribution (Hill-Winston-Boyd, 2005). While few might have expected that the students at these demanding schools would have been drawn equally from across national family incomes, the 10%/40% ratio surely demands a better understanding.

Unfortunately, ideology provides two too easy answers. One set of convictions holds that able low-income students who, in all respects, qualify for these schools are excluded by admissions policies designed to protect the children of the wealthy and well-connected from competition – these schools are "bastions of privilege."[5] A quite different ideology holds that more highly qualified students from low-income families would be welcome but they simply don't exist – that everything from inadequate nutrition to tough neighborhoods and weak families and educational systems have conspired to keep many low-income students from being able to pass a perfectly fair cut for admission to these schools.[6]

[3] Bowen (2004), Bowen, Kurzweil, and Tobin (2005) , Summers (2004), Kahlenberg (2003), Carnevale and – Rose (2003)

[4] Including Harvard, Yale, Princeton, Penn, Columbia, Dartmouth, Brown, Cornell, Duke, MIT, Stanford, Northwestern, Chicago, Georgetown, Rochester, Washington University, Rice, Johns Hopkins, Wellesley, Smith, Bryn Mawr, Barnard, Mt. Holyoke, Carleton, Oberlin, Amherst, Pomona, Trinity, Wesleyan, Williams, and Swarthmore. Three of these schools did not participate in the study.

[5] Peter Schmidt, April 16, 2004, Bowen, et al. (2005).

[6] Owings, et al (1995), Heckman (1999), Cabrera and La Nasa. (2000).

Fortunately, a third possibility can be ruled out by evidence from that same study of COFHE schools. Most of these schools practice "need blind admission with full need financial aid"[7] with the result that the price actually paid for tuition, room, board and fees by students from the poorest families (those earning under $24,000 in 2001 in our data) was sometimes less than $1,000 a year despite a mean sticker price of $33,831. On average, over all 28 schools, their net price was $7,552, before considering loans and jobs – less than the average price of a public four-year college [College Board (2004)]. So the *affordability* of these schools is not likely to be an important reason for their meager proportions of low-income students (though students' lack of knowledge of these low prices is quite likely to help explain their scarcity.[8]) At these wealthiest schools, it is generally true that a student who can get in, can afford to go.

Even without the hyperbole and conspiracy theory of a "bastions of privilege" explanation, though, there remains the possibility that at such schools, procedures exist which disadvantage highly able low-income students when they compete for admission with those from wealthier families. So two questions need investigation – are there, on the one hand, procedural biases that make it harder for low-income kids to be admitted to these places and, on the other hand, "Are they out there?" – do many high-ability, low-income students *exist* in the national population? This paper will address the second question while a paper in preparation will explore the first.

[7] So admission decisions are made without knowledge of family income and the student's price is adjusted so that tuition, loans and campus job will make the school affordable to the admitted applicant from even the lowest income family.

[8] Choy (2001), Tom Kane, (1999), "There may be substantial costs of simply learning what types of aid are available" (p.95).].

"Are they out there?" involves two further questions. What is a reasonable target for the share of low-income, high-ability students in these schools? The 10%/40% ratio may be worrisome, but what representation would not be? And second, is there a large enough number of high-ability high school students from low-income families in the US for these schools to reach a reasonable target share?

The Target: What share of these student bodies should be from low-income families?

The national population

We take as a reasonable target that the income distribution of highly able students in the national population should be mirrored in the student bodies of these highly selective schools. That, we feel, would reflect a meaningful policy of equality of opportunity. If there exists a larger share of low-income high-ability high school graduates than are found in these schools, low-income students of high-ability are under-represented. Whether that's true, of course, will depend very much on what definition of "high-ability" is chosen.

So we examine family income and test score data for ACT and SAT over all of the high school seniors who took those tests in 2003. Our assessment of "ability," then, rests on the scores achieved on either of those national tests. We use family income reported by the student test-takers. As measures of ability and family income, these data have shortcomings (discussed more fully in the Appendix), but they have the advantage of large numbers (some

2.3 million test records) and national scope, and they directly address the central question of income and ability.[9]

In Table 1, the combined national population of ACT and SAT test-takers is reported by minimum score and by income divided into US Census family income quintiles.[10] ACT scores are expressed as SAT-equivalents and merged. Under each test score, the first row of the table describes the number of students in the national population who achieved that score *or higher* in each income quintile. The second row shows their distribution (as a percent of those who reported income). So, for instance, the top rows of Table 1 show the number of students scoring 1600 and their percentage distribution across family incomes. The next two rows describe those scoring 1520 or above: 193 or 2.6% of the 7,425 students who scored 1520 or above[11] came from the bottom income quintile, 598 or 8.1% came from the second income quintile, 1,052 or 14.2% from the third, and so on. The next rows use a minimum score of 1420 to define ability, and so on down to a combined SAT of 400 or above which, of course, includes the whole population of test-takers. This gives a convenient way to represent the income distribution of students in the population as they are defined by different minimal levels of ability.

[Table 1]

[9] What's more, the use of test scores alone to measure student ability gains plausibility from a 1995 NCES study that described a more complex set of five criteria as necessary for admission to a highly selective college and found that test score was among the most discriminating of single criteria, eliminating 80% of the population even when they used an 1100 minimum SAT level. [Owings, et. al., 1995]

[10] For details on the Census boundaries and extrapolations for quintile medians, see the Appendix to Hill-Winston-Boyd.

[11] See the Appendix for a discussion of non-reporting of incomes.

It is important that the issue embedded in these numbers is different from that of much of the literature which has looked at the share of high school graduates who meet various college admission criteria. So the NCES study reported, for instance, that of all "college bound" high school graduates in 1992, only 5.9% satisfied all five of the criteria they identified as needed for admission to a highly selective college[12] and they examined the effect on the number who passed muster of relaxing each of those criteria.

Ours is a different question. We ask, "Of those who meet various minimum SAT-ACT criteria – various potential specifications of high-ability – how many of them come from families in each of the five income quintiles?" Table 1, then, describes the income distribution of high-ability students, variously defined, in the national population while theirs describe how many students in the college going population <u>are</u> highly able.[13]

Two Appendix tables report ACT and SAT data separately; we base our conclusions on the combined population, but since we have no way to avoid the double counting of those who took both tests, we confirm our conclusions using the two tests separately. Within each test population, any test-taker appears only once. And, as noted, the Appendix includes a discussion of the shortcomings of these data in addressing the "Are They Out There?" question and the effects of those shortcomings.

[12] A senior high school cumulative GPA of 3.5; SAT (or ACT equivalent) of 1100; four English Courses, three each in math, sciences and social science and two in foreign language; positive teacher evaluations; and evidence of engagement in extracurricular activities. (Owings, et. al., 1995).

[13] It is reassuring that when that 1995 NCES study (Owings, et. al. (1995) looked only at those who met all five of their criteria for selectivity and grouped them by family income (SES), they reported results that were quite consistent with those in Table 1 – 10.4% of highest ability students came from families with the lowest SES.

The series of Figures 1-a through 1-m picture the information in the rows of Table 1 – the national distribution of test takers over family incomes for those who scored 1600, those who scored 1520 or above, 1420 or above, and so on down to 400 and above. So, together, these graphs describe the income distribution of that part of the national student population. that meets and exceeds alternative definitions of ability.

[Figures 1-a to 1-m.]

It's clear from the figures that with declining minimal test scores the changing shape of the income distributions of the populations they define describes the pattern familiar in national data – as test score minima fall, the corresponding income distribution flattens until, for those scoring 400 or above in Figure 1-m (the whole of the population), the distribution is roughly equal across income quintiles.[14] It is clear, too, that the answer to "Are They Out There?" is very sensitive to the choice of a definition of "high-ability," so we report alternative measures of high-ability.

Low-income students in the COFHE schools

Table 2 is taken from the study of pricing at COFHE schools that triggered the question about whether there are more high-ability, low-income students in the population than are found in these schools. With the same structure as Table 1, it reports the number of students in the COFHE undergraduate population and their percentage distribution over the five income quintiles. With 5% from the first quintile (under $24,000) and 5% from the

[14] We would not expect it to be equal across the national quintiles because the test-takers are not representative of the national family population.

second ($24,001 to $41,000) we get the fact we started with: that 10% of these students come from the bottom 40% of the US family income distribution.

[Table 2]

Comparing the data in Tables 2 and 1, we can see what the COFHE schools would have to do to mirror the income distribution of the national population of high-ability students. Defining high-ability with the rather ambitious specification of minimum ability at or above an SAT of 1420, for instance, Table 1 indicates that 12.8% of those scoring that high or higher in the national population come from families in the bottom two income quintiles. The COFHE schools, then, could increase their share of low-income students by nearly 30% while maintaining that 1420 minimum standard of ability. If a score of 1300 or above were considered adequate, 16% of those who qualify would need to come from the bottom two quintiles – an increase in share by more than half. (The NCES ability cutoff of 1100 [Owings, et.al. 1995] would see more than 21% from low-income families and Carnevale and Rose's definition of high-ability as 900 minimum score [2004] would have nearly 30% of the COFHE students from the poorest 40% of the families.)

The lower the ability threshold, predictably, the larger the share of students who would come from the bottom two quintiles. So what are reasonable targets for these highly selective schools? Some data provide a good sense of an answer with information on the inter-quartile range of their test scores. In the average COFHE school, 25% of their students score under 1353 and 25% score over 1546. The lowest school score at the 25th percentile is 1160 and the highest is 1400. At the other end, the lowest 75th percentile score is 1375 and the highest is

1580. So it seems reasonable to focus on the income distributions of the six populations in Table 1 that correspond to ability levels defined by minimal scores from 1110 to 1600.

A more complex picture comes from the series of Figures 2 where the national information on income distribution by minimum scores – the bars from Figures 1 – are repeated and contrasted to a similarly defined picture of the income distribution at the COFHE schools. We have little information on test scores for the COFHE students – only their inter-quartile range and distribution over incomes in the aggregate – so from one panel to the next, the bars describing the schools' populations are the same – in each panel, the COFHE distribution is compared to that of a national population defined by alternative levels of ability, from 1110 to 1600.

[Figures 2]

It's apparent from the first few figures that the COFHE schools, collectively, do very well by the lowest income, highest ability students. Of those from the bottom quintile who score 1600, a larger share is found in the COFHE schools than in the national population.[15] Moving from the first to the second income quintile of students scoring that high, however, COFHE's relative share of low-income students drops and within the rest of the distribution (the top 60%), it is the middle and upper middle income students who are markedly under-represented while those from the highest income families are over-represented at these schools.

[15] This, of course, is consistent with the evidence from Hoxby (1997) and Frank (2001) of the concentration of the nation's best students at a few colleges.

As the measure of high-ability is moved down the SAT scale in Figure 2-b and beyond, the under-representation of middle income students comes to include students increasingly from lower-middle income families and, indeed, even the under-representation of students from the bottom two quintiles, relative to the national population, appears to be due primarily to that second quintile – the lowest income students scoring, for instance, 1520 or above, are over-represented at these schools but not those at higher incomes. Defining high-ability with a minimum score of 1420, the lowest incomes are only slightly over-represented while the next three quintiles – from $24,000 to $92,000 – are under-represented. For scores 1300 and above, the scarcity of middle-income students becomes more pronounced and here Figure 2-d shows that all but the highest income students at that ability level are under-represented.[16]

The Numbers: Are They Out There?

While the appropriate description of the issue – and the implied policy target – appears to rest on alternative targets that reflect alternative ability levels through which these schools would mirror the population share of high-ability low-income students, there remains an important question of numbers. In both directions. We've said, for instance, that if high-ability is defined as an SAT score of 1420 or above, the COFHE schools should move, collectively, from having 10% to nearly 13% of their students come from the bottom two family income quintiles – they should increase their low-income population by roughly 30%. But can they? What does that mean both for them and for the national availability of high-ability low-income students in the population? If we change the "high-ability" definition to

[16] These figures, of course, raise the question of whether the target of the COFHE schools should be a representative distribution of income over all income levels or whether there is a special obligation, with a goal of equality of opportunity, to those high-ability students from the lowest income families.)

1300-or-above or to 1220-or-above, what does that imply about targets, numbers, and the availability of those students?

A good first cut at "are the numbers out there?" comes from adding up the figures in the low and lower-middle income cells of Table 1 for each specific ability level – the supply or availability of such students in the national population – and comparing that sum with the number necessary to achieve the target implied by that ability level – the COFHE schools' demand for such students. So, again, if 1420 were taken as the minimum ability level, the present 10% in COFHE schools from the bottom two income quintiles would have to be increased to 12.8% – instead of the 2,750 low-income students now matriculating per year,[17] there would have to be 3,520. That's the schools' demand. On the supply side, Table 1 indicates that at 1420 and above, there are, nationally, 4,276 students in those bottom two income quintiles. So meeting that target is not impossible, but it's tight: nearly 85% of the low-income, high-ability students in the US would have to go to one of these COFHE schools in order for them to mirror national population shares under that definition of high-ability.[18] If the high-ability definition were reduced to a minimum score of 1300, the enrollment target would become 16% which means that 4,400 would have to be matriculated each year from the low-income population of 19,959 who score 1300 or above: 22% of the national high-ability, low-income population would be in COFHE schools. At 1220 or above, the schools' target would be 18.2% or 5,005 low-income students per year from a national population of nearly 44,000. As the definition of high-ability is relaxed, the target enrollment increases, but not

[17] The schools in our data have a total undergraduate enrollment of approximately 110,000 which implies that 27,500 students are being matriculated each year. With 10% of them from the lowest income families, 2,750 low-income students matriculate yearly. A 28% increase, then, would see an additional 770 low-income students each year.

[18] For comparison, the COFHE schools enroll approximately 19,000 students a year from the fifth quintile, while 15,238 students from the fifth quintile score 1420 or above.

nearly as fast as does the population of low-income students available to meet that target. These facts are summarized in Table 3 for ability levels of 1110 and above to 1600.

[Table 3]

So it appears that for reasonable ability levels, these schools under-represent low-income students and, for those ability levels, the students are out there. Some shortcomings with the data, though, make these conclusions a bit too easy.

A large fraction of the test-takers – ranging from 35% to over 50% in these score ranges – simply don't report family income so they could not be included in the numbers or percentage distributions of Table 1. Since studies of the distribution of non-reporting over family incomes (see the Appendix) suggest that low-income students are more likely not to report family income than are high income students, non-reporting most likely leads to understatement of the available population share of high-ability low-income students and of any associated benchmark.

In the other direction, we know that we double count those who take both the SAT and the ACT tests, so we know that we overstate the numbers in the data of Table 1 that combines them – and we know that we have no hint of the degree of overstatement. But, like the non-reporting of income, it seems quite likely that high-income students more frequently take both tests than do low-income students. While we've tested our conclusions about the distribution of students over incomes by using SAT and ACT data separately (Appendix tables 1 and 2), we can't do that to verify our counts of available populations by score. So double counting of

those who took both ACT and SAT will have overstated the number out there, but that overstatement would appear to affect mostly the high income students. [19]

Conclusion:

There are two components of "are they out there?" One looks at plausible targets for expansion of low-income enrollment at these highly selective schools by comparing the income distribution of their present students with that of the national population of high-ability, low-income students. Implicit is that these schools should aim to have a population of such students that mirrors the national population and the data suggest that there is considerable room for them to do that – to provide these students with greater equality of opportunity. The other component asks if, given these targets, the numbers are out there – are there enough low-income students of high-ability in the US realistically to allow these schools to mirror the national low-income distribution of the highly able students? The answer to that is "Probably, but it depends" on the definition of high-ability adopted – as that definition is relaxed, the population of available low-income students increases quickly. At ability levels that look quite reasonable in light of current COFHE scores – say 1220 to 1420 – there appear to be plenty of low-income students.

[19]Redoing Table 3, underestimating supply by using only those who took the ACT – thereby precluding double counting – would show that there are 2,180 low-income students scoring 1420 or above (against a demand of 3,520) while 11,000 who score 1300 or above are from low-income families (against a demand of 4,400).

Table 1
The Distribution of Students over Family Income by Ability Level
National SAT and ACT Test-taking Population Combined
(2003)

SAT Equivalent Score	Family Income						
Income	**Lowest**	**Lower Middle**	**Middle**	**Upper Middle**	**High**	**Total Reporting Income**	**No Income Report**
Lower Bound	--	$24,001	$41,001	$61,379	$91,701		
Quintile Median	$15,347	$32,416	$50,890	$74,418	$113,689		
1600	7	30	48	112	252	449	506
Percent	1.6%	6.7%	10.7%	24.9%	56.1%	100%	
1520 & above	193	598	1,052	1,871	3,711	7,425	5,116
Percent	2.6%	8.1%	14.2%	25.2%	50.0%	100%	
1420 & above	1,229	3,047	5,363	8,406	15,288	33,333	20,776
Percent	3.7%	9.1%	16.1%	25.2%	45.9%	100%	
1300 & above	5,982	13,977	23,318	32,912	48,747	124,936	70,334
Percent	4.8%	11.2%	18.7%	26.3%	39.0%	100%	
1220 & above	13,360	30,238	47,683	63,113	85,448	239,842	127,219
Percent	5.6%	12.6%	19.9%	26.3%	35.6%	100%	
1110 & above	36,304	72,706	104,950	128,841	152,152	494,953	238,079
Percent	7.3%	14.7%	21.2%	26.0%	30.7%	100%	
1030 & above	62,404	117,124	158,043	184,752	198,566	720,889	329,027
Percent	8.7%	16.2%	21.9%	25.6%	27.5%	100%	
910 & above	122,412	199,916	245,299	266,401	257,655	1,091,683	464,440
Percent	11.2%	18.3%	22.5%	24.4%	23.6%	100%	
830 & above	173,758	256,333	296,051	307,457	283,905	1,317,504	540,701
Percent	13.2%	19.5%	22.5%	23.3%	21.5%	100%	
740 & above	227,465	303,930	330,957	334,082	297,906	1,494,340	599,439
Percent	15.2%	20.3%	22.1%	22.4%	19.9%	100%	
620 & above	270,223	332,459	348,452	345,527	303,641	1,600,302	638,845
Percent	16.9%	20.8%	21.8%	21.6%	19.0%	100%	
500 & above	286,576	341,063	353,194	348,324	305,050	1,634,207	653,726
Percent	17.5%	20.9%	21.6%	21.3%	18.7%	100%	
400 & above	289,061	342,113	353,654	348,653	305,207	1,638,688	655,841
Percent	17.6%	20.9%	21.6%	21.3%	18.6%	100%	

Table 2
The Distribution of Students over Family Income
28 Highly Selective Private Schools
(2001-2002)

	Family Income					Total Enrollment
Income	**Lowest**	**Lower Middle**	**Middle**	**Upper Middle**	**High**	
Lower Bound	-	$24,001	$41,001	$61,379	$91,701	
Quintile Median	$15,347	$32,416	$50,890	$74,418	$113,689	
Number of Students						
COFHE Schools	5,086	5,956	8,053	12,086	75,803	108,721
Coed Colleges	698	958	1,242	1,951	10,501	15,471
Women's Colleges	532	641	752	962	5,515	8,620
Ivy League Universities	2,079	2,290	3,130	4,747	32,870	45,609
Non-Ivy Universities	1,777	2,067	2,929	4,426	26,918	39,022
Percent of Total Enrollment						
COFHE Schools	5%	5%	7%	11%	70%	100%
Coed Colleges	5%	6%	8%	13%	68%	100%
Women's Colleges	6%	7%	9%	11%	64%	100%
Ivy League Universities	5%	5%	7%	10%	72%	100%
Non-Ivy Universities	5%	5%	8%	11%	69%	100%

Table 3
The Availability of Low Income, High Ability Students
To Meet Yearly Enrollment Targets

Ability	Low Income Target Enrollment: Schools' Demand					Available Numbers: National Supply	
SAT Equivalent Score	Target	Low Income Enrollment Increment Needed to Reach Target			Total Yearly Low Income Demand		
		Percent of Student Body	Percent of Low Income Students	Number of Additional Low Income Students		Q1 + Q2	Maximum Including Non-reports*
1600	8.3%	-1.7%	-17.0%	-468	2,283	37	543
1520 & above	10.7%	0.7%	7.0%	193	2,943	791	5,907
1420 & above	12.8%	2.8%	28.0%	770	3,520	4,276	25,052
1300 & above	16.0%	6.0%	60.0%	1650	4,400	19,959	90,293
1220 & above	18.2%	8.2%	82.0%	2255	5,005	43,598	170,817
1110 & above	22.0%	12.0%	120.0%	3300	6,050	109,010	347,089

The COFHE undergraduate student population is approximately 110,000, of whom 10% are from income quintiles 1 and 2 (table above). With 110,000/4=27,500 added each year, 2750 low income students each year would represent a steady-state.

* If all income non-reports were in the bottom two quintiles.

Table A-1
The Distribution of Students over Family Income by Ability Level
National SAT Test-taking Population Only

SAT Equivalent Score / Income	Family Income: Lowest	Lower Middle	Middle	Upper Middle	High	Total Reporting Income	No Income Report
Lower Bound	--	$24,001	$41,001	$61,379	$91,701		
Quintile Median	$15,347	$32,416	$50,890	$74,418	$113,689		
1600	0	20	20	70	190	300	460
Percent	0.0%	6.7%	6.7%	23.3%	63.3%	100%	
1520 & above	80	300	360	830	2,290	3,860	3,820
Percent	2.1%	7.8%	9.3%	21.5%	59.3%	100%	
1420 & above	630	1,430	2,310	4,130	9,830	18,330	15,560
Percent	3.4%	7.8%	12.6%	22.5%	53.6%	100%	
1300 & above	2,790	6,050	9,860	15,990	29,210	63,900	50,080
Percent	4.4%	9.5%	15.4%	25.0%	45.7%	100%	
1220 & above	5,950	12,870	19,870	29,930	50,200	118,820	88,590
Percent	5.0%	10.8%	16.7%	25.2%	42.2%	100%	
1110 & above	15,610	29,550	41,770	59,450	85,990	232,370	159,560
Percent	6.7%	12.7%	18.0%	25.6%	37.0%	100%	
1030 & above	26,510	47,660	62,970	85,250	110,290	332,680	217,380
Percent	8.0%	14.3%	18.9%	25.6%	33.2%	100%	
910 & above	51,300	79,690	96,720	121,440	140,010	489,160	297,840
Percent	10.5%	16.3%	19.8%	24.8%	28.6%	100%	
830 & above	71,850	100,580	116,370	139,300	153,060	581,160	339,340
Percent	12.4%	17.3%	20.0%	24.0%	26.3%	100%	
740 & above	92,510	119,080	130,280	151,750	159,960	653,580	370,160
Percent	14.2%	18.2%	19.9%	23.2%	24.5%	100%	
620 & above	110,890	131,080	138,050	157,380	162,860	700,260	390,630
Percent	15.8%	18.7%	19.7%	22.5%	23.3%	100%	
500 & above	119,080	134,920	140,480	158,970	163,620	717,070	397,930
Percent	16.6%	18.8%	19.6%	22.2%	22.8%	100%	
400 & above	120,830	135,580	140,770	159,230	163,740	720,150	399,320
Percent	16.8%	18.8%	19.5%	22.1%	22.7%	100%	

Table A-2
The Distribution of Students over Family Income by Ability Level
National ACT Test-taking Population Only

SAT Equivalent Score / Income	Family Income: Lowest	Lower Middle	Middle	Upper Middle	High	Total Reporting Income	No Income Report
Lower Bound	--	$24,001	$41,001	$61,379	$91,701		
Quintile Median	$15,347	$32,416	$50,890	$74,418	$113,689		
1600	7	10	28	42	62	149	46
Percent	4.7%	6.7%	18.8%	28.2%	41.6%	100%	
1520 & Above	113	298	692	1,041	1,421	3,565	1,296
Percent	3.2%	8.4%	19.4%	29.2%	39.9%	100%	
1420 & Above	599	1,617	3,053	4,276	5,458	15,003	5,216
Percent	4.0%	10.8%	20.3%	28.5%	36.4%	100%	
1300 & Above	3,192	7,927	13,458	16,922	19,537	61,036	20,254
Percent	5.2%	13.0%	22.0%	27.7%	32.0%	100%	
1220 & Above	7,410	17,368	27,813	33,183	35,248	121,022	38,629
Percent	6.1%	14.4%	23.0%	27.4%	29.1%	100%	
1110 & Above	20,694	43,156	63,180	69,391	66,162	262,583	78,519
Percent	7.9%	16.4%	24.1%	26.4%	25.2%	100%	
1030 & Above	35,894	69,464	95,073	99,502	88,276	388,209	111,647
Percent	9.2%	17.9%	24.5%	25.6%	22.7%	100%	
910 & Above	71,112	120,226	148,579	144,961	117,645	602,523	166,600
Percent	11.8%	20.0%	24.7%	24.1%	19.5%	100%	
830 & Above	101,908	155,753	179,681	168,157	130,845	736,344	201,361
Percent	13.8%	21.2%	24.4%	22.8%	17.8%	100%	
740 & Above	134,955	184,850	200,677	182,332	137,946	840,760	229,279
Percent	16.1%	22.0%	23.9%	21.7%	16.4%	100%	
620 & Above	159,333	201,379	210,402	188,147	140,781	900,042	248,215
Percent	17.7%	22.4%	23.4%	20.9%	15.6%	100%	
500 & Above	167,496	206,143	212,714	189,354	141,430	917,137	255,796
Percent	18.3%	22.5%	23.2%	20.6%	15.4%	100%	
400 & Above	168,231	206,533	212,884	189,423	141,467	918,538	256,521
Percent	18.3%	22.5%	23.2%	20.6%	15.4%	100%	

Figure 1-a.
The Distribution over Family Income,
Those with an SAT Equivalent Score of
1600

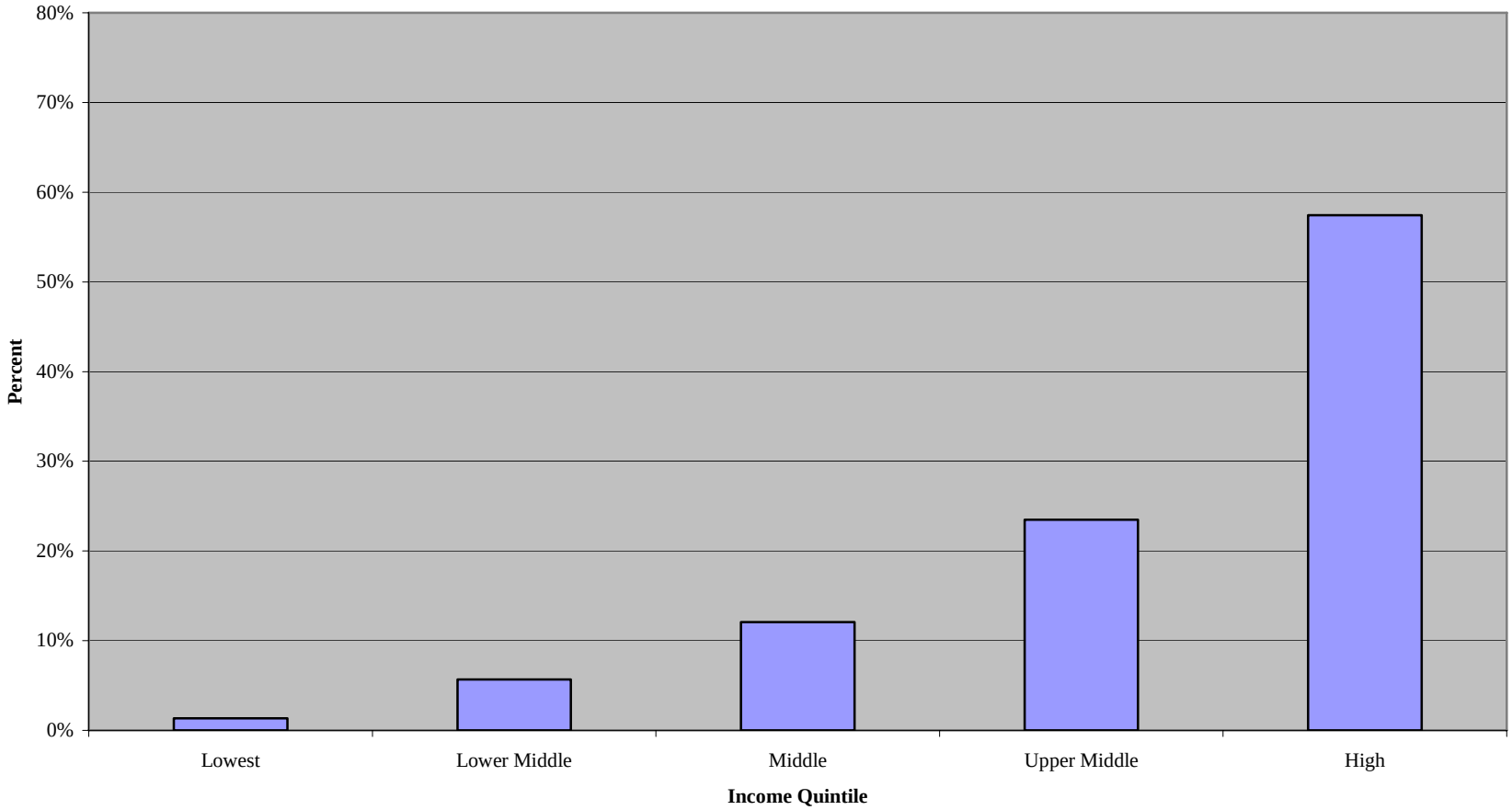


Figure 1-b.
Those with SAT Equivalent Scores of
1520 & above

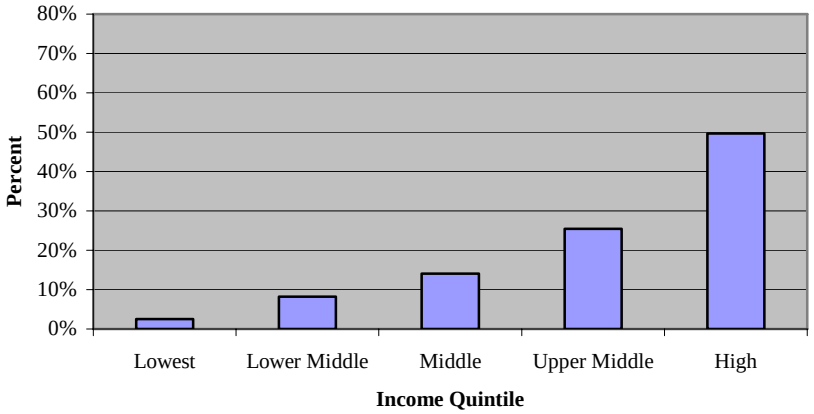


Figure 1-c.
Those with SAT Equivalent Scores of
1420 & above

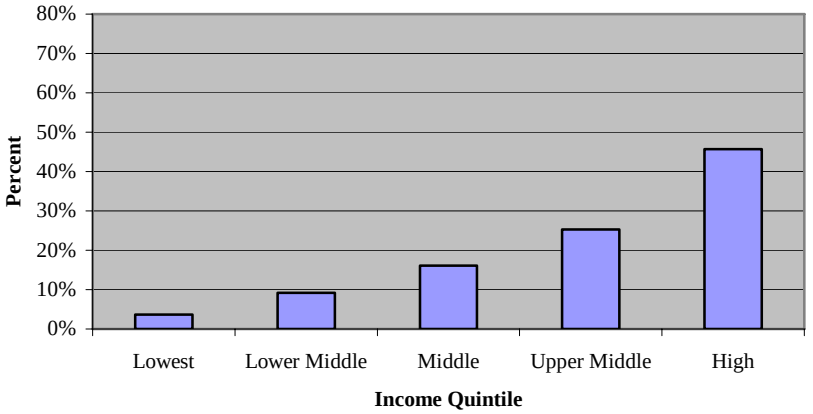


Figure 1-d.
Those with SAT Equivalent Scores of
1300 & above

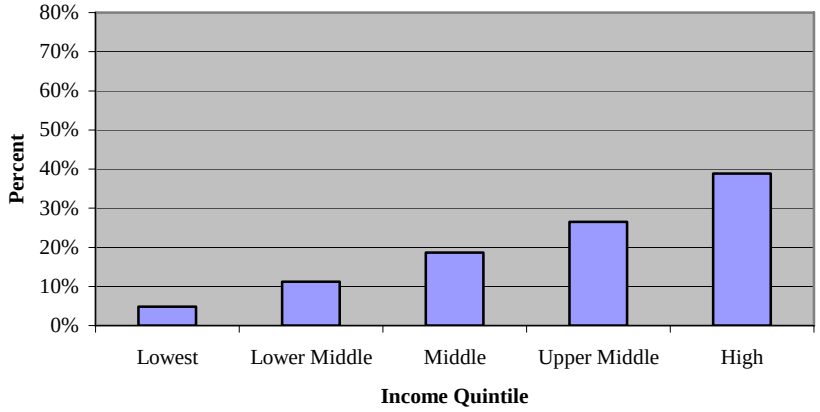


Figure 1-e.
Those with SAT Equivalent Scores of
1220 & above

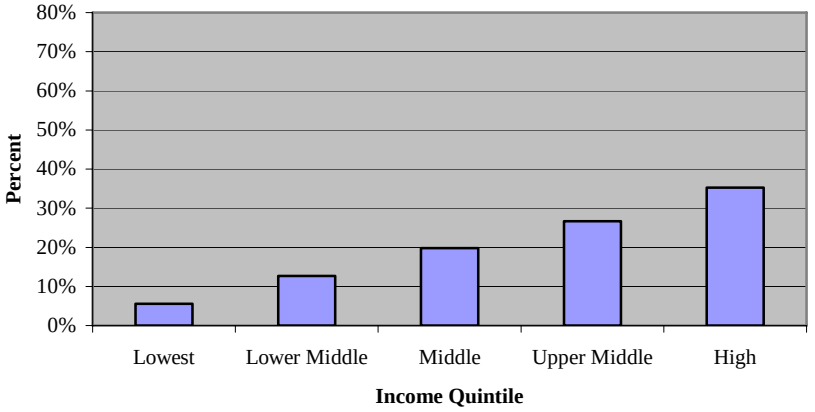


Figure 1-f.
Those with SAT Equivalent Scores of
1110 & above

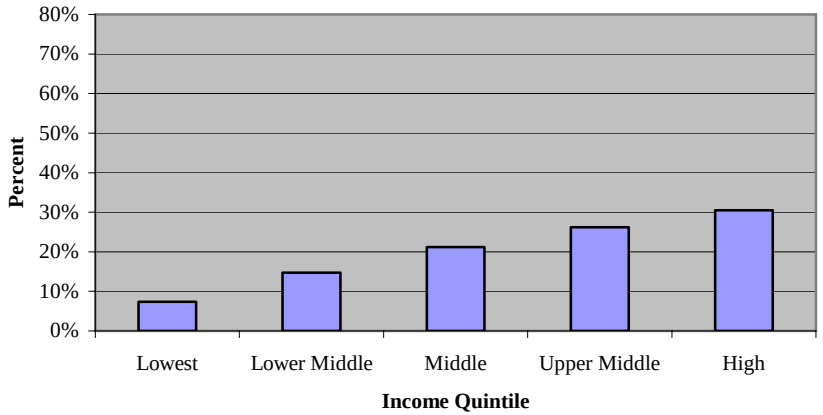


Figure 1-g.
Those with SAT Equivalent Scores of
1030 & above

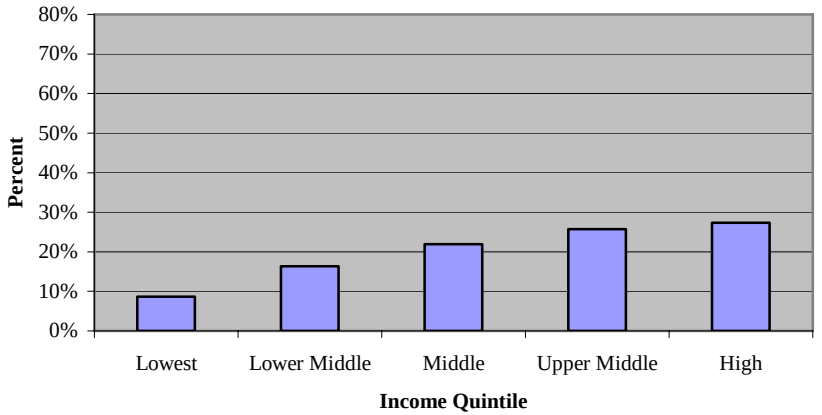


Figure 1-h.
Those with SAT Equivalent Scores of
910 & above

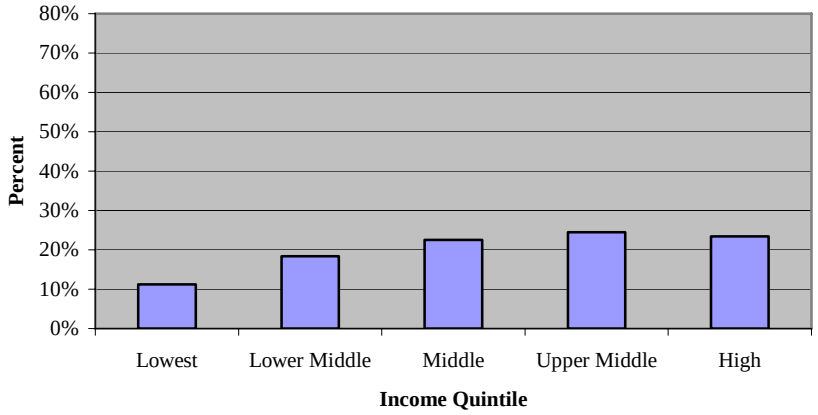


Figure 1-i.
Those with SAT Equivalent Scores of
830 & above

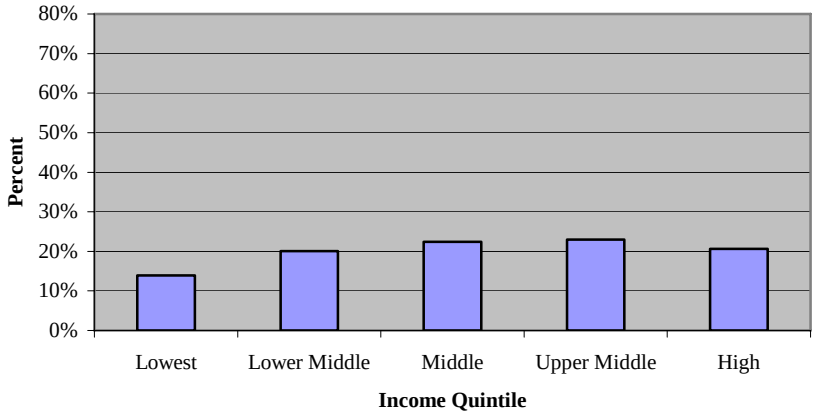


Figure 1-j.
Those with SAT Equivalent Scores of
740 & above

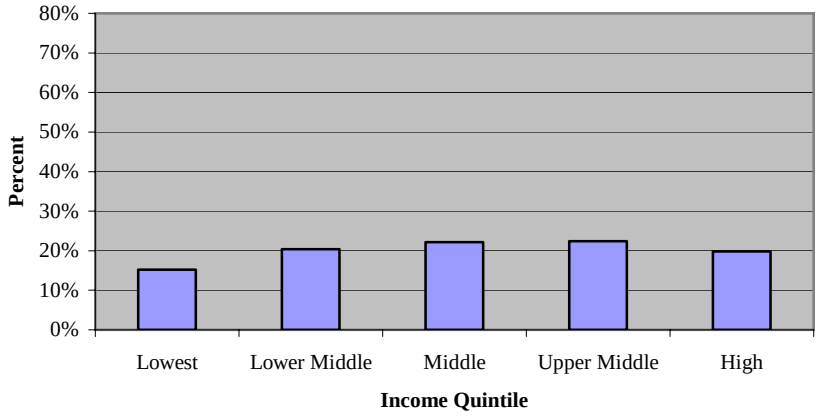


Figure 1-k.
Those with SAT Equivalent Scores of
620 & above

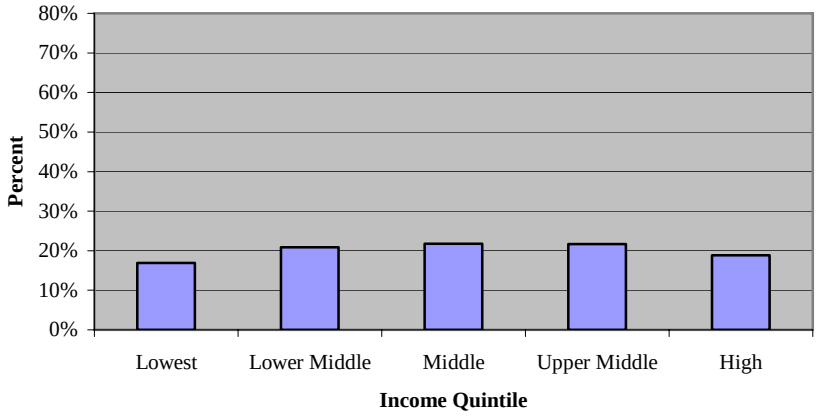


Figure 1-l.
Those with SAT Equivalent Scores of
500 & above

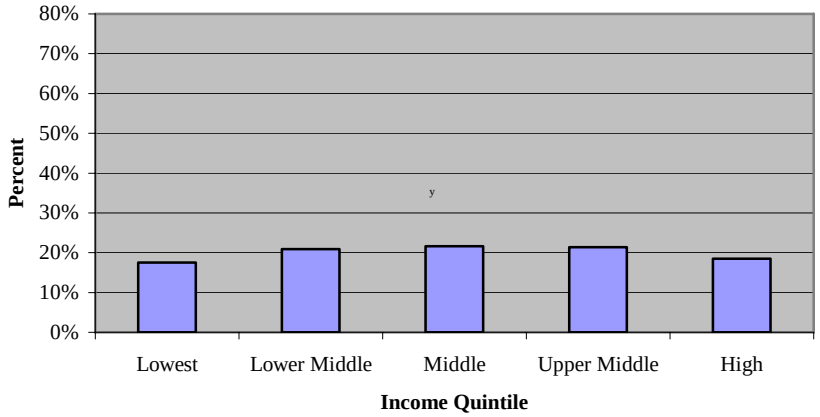
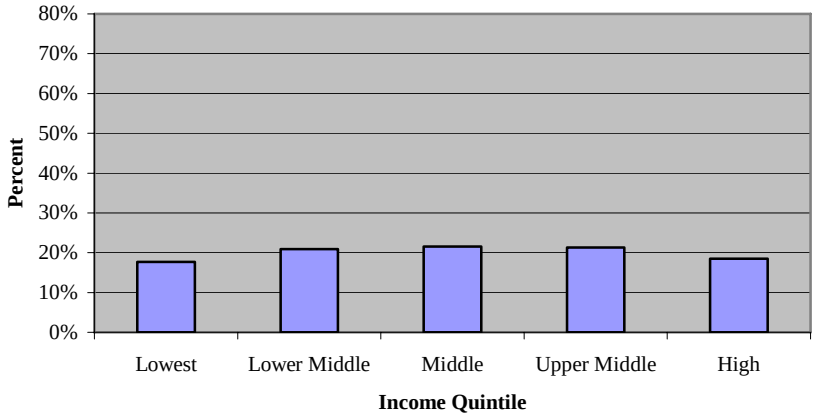


Figure 1-m.
Those with SAT Equivalent Scores of
400 & above



COFHE, Compared to the National Population

Figure 2-a.
Those with an SAT Equivalent Score of
1600

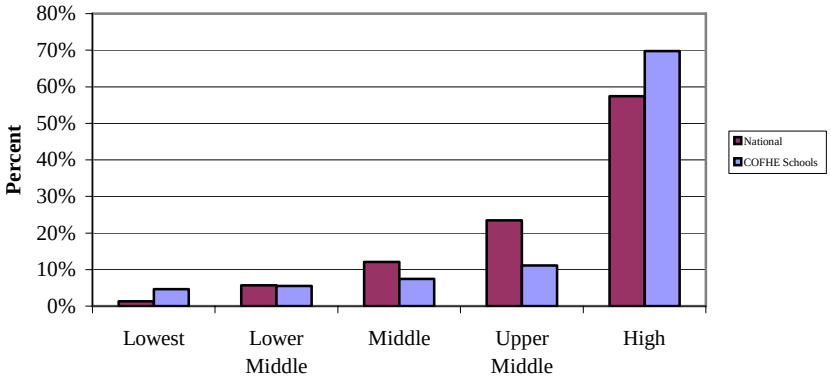


Figure 2-b.
Those with SAT Equivalent Scores of
1520 & above

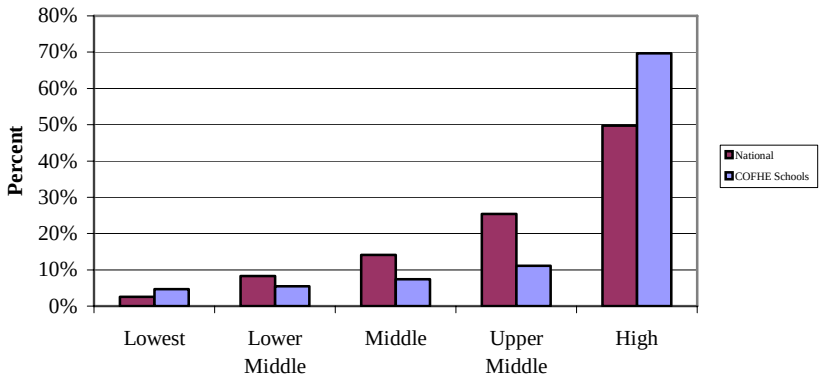


Figure 2-c.
Those with SAT Equivalent Scores of
1420 & above

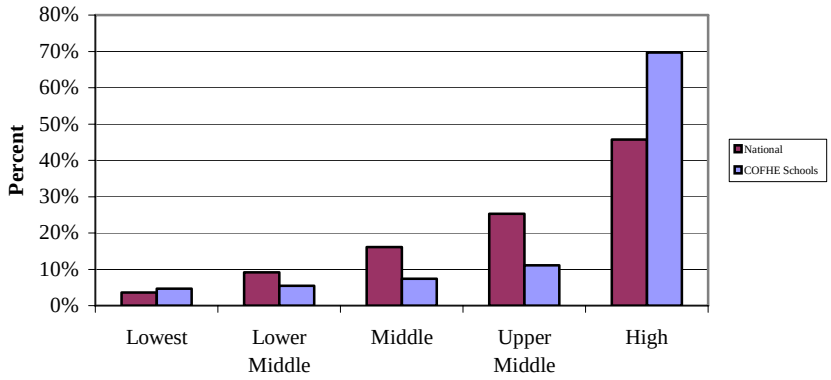


Figure 2-d.
Those with SAT Equivalent Scores of
1300 & above

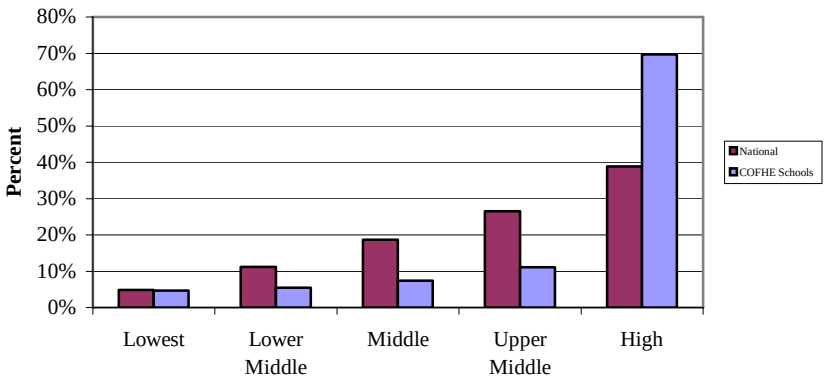


Figure 2-e.
Those with SAT Equivalent Scores of
1220 & above

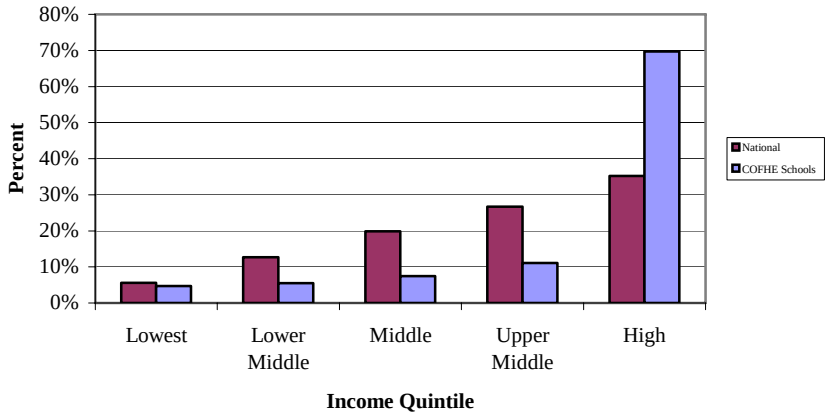
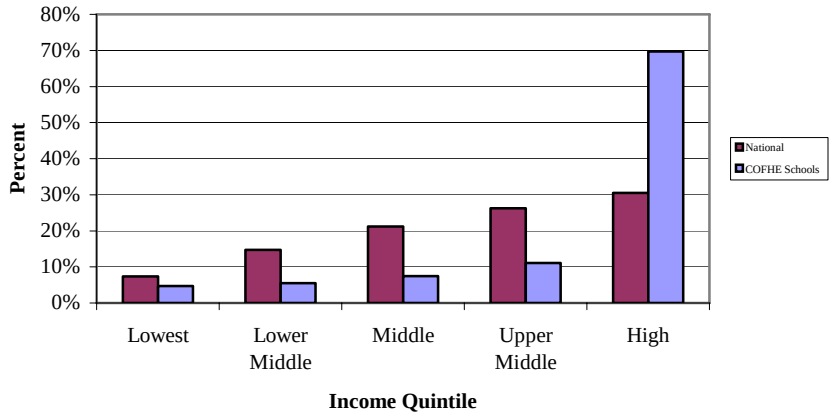


Figure 2-f.
Those with SAT Equivalent Scores of
1110 & above



APPENDIX

The Quality of the Data

We rely on test scores and students' self reported incomes to measure ability and family incomes. There are good reasons to be uncomfortable with these and to hope that both the sheer size of the population and the direction of potential distortions give us persuasive results. So it becomes important that these data are being used to describe the share of high-ability low-income students in the national test-taking population and compare that with their share in the student bodies of this set of highly selective private colleges and universities. As noted in the text, since we conclude that a larger share of the high-ability kids comes from low-income families in the US population than in the COFHE schools, data problems will weaken that conclusion only if they might lead to overstatement of the population share of those high-ability, low-income students.

Student Ability Data

That said, ideally we would judge student ability on the kind of information available to colleges in making their admission decisions – GPA, class rank, high school courses, extracurricular activities, test scores, etc. across all test-takers. The 1995 NCES study (Owings, et al, 1995) used NELS data to approximate that. Instead, we rely solely on reported test scores, giving a thin measure of ability, but in a very large national database (more than two million test-takers) that also reports variables including each individual's race, gender and, especially usefully, geographic location. Since the issue in this paper is ability relative to family income, the imperfections in our test-score data that raise legitimate worries are those due to (a) multiple test-taking which has been shown to

raise scores and (b) test preparation courses that do the same thing. For our purposes, then, it is reassuring that while both of these appear to be income-sensitive (Vigdor and Clotfelter, 2003)**,** raising the relative test scores of high income students, they therefore tend to understate the share of low-income students in the population who would, with the same advantages, show high ability. So a measure of "true" ability (unaffected by repetitive testing or test prep) would show a larger, not a smaller, share of high-ability low-income students in the test-taking population.

A different problem is introduced by our combining the populations of ACT and SAT test-takers: double counting is eliminated within each test, but we have no way of knowing how many or which students in our data are counted twice because they've taken both tests. While individual admissions departments can and do scan their search data to identify such duplication, we have no way to do that. Therefore, while we've reported the results for the combined populations – expressing ACT scores in SAT equivalents and adding the populations together – we have confirmed all important findings by running the two data-sets separately.

Family Income Data

The use of students' self-reported family incomes raises questions of accuracy, on the one hand, and of the effect of a large number of non-reporting students, on the other. More than fifty percent of those scoring an SAT equivalent of 1600 on ACT and SAT combined don't report family income at all. Lower scores showed a smaller incidence of non-reporting.

The Accuracy of Reported Incomes

Inaccurate income reporting, per se, would simply add noise to the analysis but if high- or low-income students report inaccurately in different ways or different degrees, our descriptions of high-ability income distribution might be affected. As noted, however, only if lower-income students were more prone to understate family incomes, would their share in the high-ability population be overstated in our data.

There is clear recognition in the literature that self-reported incomes can be unreliable (see Terenzini, *et. al.* (2001), *inter alia*) but there is consensus on neither methodology of judging their accuracy nor the shape of any inaccuracies. Studies variously compare family income reports by high school students (Kayser and Summers, (1973) or community college or public university freshmen (Romano and Moreno (1994); Smith and McCann (1998)) or adults (Moore, Stinson and Welniak (2000)) with federal income tax returns (Romano and Moreno (1994); Smith and McCann (1998)) or parental income reports (Fetters, Stowe and Owings (1984); Kayser and Summers (1973)). Most conclusions are based on comparisons within families but one rests on aggregated state data (Card and Payne (2002)). There is little agreement on how to measure "accuracy", with some studies appearing to generate their conclusions from their methodology.[20] Finally, some studies considered a number of possible influences on accuracy including age, gender, education and family socioeconomic status and some

[20] As in Romano and Moreno's (1994) use of eleven $3,000-wide income brackets, that led them to conclude that low-income subjects report income more accurately, neglecting the fact that a much larger percentage error was needed to miss a low-income bracket than a high-income one.

embed the issue of income accuracy in the question of a broader socioeconomic index. Some studies suggest regression to the mean, concluding that low-income subjects tend to report too much income and high-income subjects too little (Smith and McCann (1998)); some report general overstatement of income (Kayser and Summers (1973)); some report greater accuracy by low-income students (Romano and Moreno (1994)).

Our conclusion is that no clear picture has emerged from earlier studies that would lead us to believe that biases in the family incomes self-reported by SAT and ACT test-takers inflate the apparent share of high-ability students from low-income families in the national population. Williams' data, described below, reinforces that conclusion: they not only allowed us to look at both the accuracy of self-reported income and the incomes of those who don't self-report income but to do it with a population that represents the high-ability students in which we are interested in this study.

Failure to report family income

The literature on income surveys understandably provides only weak reassurance about the incomes of those who don't report income. Our concern, of course, is that systematic bias in non-reporting by income level might be inflating the apparent share of high-ability low-income students – if most non-reporting is done by high-income students then the share of high-ability low-income students in the population would clearly be overstated. Again, the evidence is mixed. Hasseldenz (1976) finds (in a mail survey of Kentucky adults) that non-respondents have lower incomes (by about 20% on average) than those who responded. (Kentucky Department of Revenue incomes were used as

"truth.") Another study (Ross and Reynolds (1996)) took a more indirect tack, finding that those lacking household and socioeconomic power and trust were less likely to respond to questions about income in a national telephone survey. These are both pretty far away from the high school test-takers whose income estimates we use, but the suggestion is that low-income students are less, not more, likely to report income. Griffin (2003) found both regional and urban-rural differences in response rates, but it is hard to know what to do with that in the present context.

The Evidence from Williams' data

We're lucky to have a highly relevant, albeit small, population with which to examine both questions of the accuracy of students' family income reports and the incomes of those who don't report incomes. We used 6 years of data from those 1440 Williams' students who both filled out the American Freshman Survey with its self-reported family income and subsequently applied for financial aid, thereby providing an IRS Form 1040 (which we take as as close to the truth about family income as we're likely to come). While the fact that all of these were applicants for financial aid might appear to draw only from a low-income population, the considerable cost of going to Williams (currently a sticker price of $38,000) means that the income range was not as truncated as one might expect.[21]

The two relevant facts that emerged from these data were that low-income students are quite accurate in their estimates of family income and that it is they, rather than higher income students, who are more likely not to report income at all. If these

[21] Indeed, in our population, a number of students had family incomes above $100,000.

same behaviors are reflected in the incomes that high-ability students report to ACT and SAT, distortions in the self-reported income data aren't inflating our measure of the share of high-ability, low-income students in the national population. We will be more confident, of course, when we've finished our ongoing analysis of a larger number of students, but in the meantime, these results are both directly to the point and reassuring. So while it remains a possibility, there is no convincing empirical evidence that distortions in self-reported income would cast serious doubt on our analysis.

SAT and ACT results separately

The following tables report the SAT and ACT data separately.

[TABLE A-1]

[TABLE A-2]

It's clear that the conclusions of the text remain. Though the ACT data suggest that there are more low-income, high-ability students in the general population than do the SAT data, both indicate that there is a higher proportion "out there" than the ten percent in the COFHE schools. At a minimum score of 1420, for instance, the SAT population shows that 11.2% are from the bottom two quintiles while the ACT population shows 14.8%. Only for those scoring 1520 to 1600 on the SAT is the share of those from the bottom two quintiles under COFHE's 10% -- a very demanding definition of 'high-ability.'

Note that those taking the ACT are in general of lower income than those taking the SATs. So of all SAT takers, 35.8% are in the bottom two quintiles while among ACT takers, 40.8% are in those quintiles. Looking at the other end of the distribution, 44.6% of those taking the SAT are in the top two income quintiles while only 36% of those taking the ACT are. Put the other way, the ratio of high- to low-income test takers is 1.25 for the SATs and 0.88 for ACT. Restricting attention to those who score 1110 and above gives the same picture: in the SAT population, 19.6% are from the low-income quintiles while 62.5% are from the top quintiles; in the ACT population, those shares are 24.3% and 51.7% respectively and the ratios of high- to low-income test-takers at this ability level are 3.2 for the SAT and 2.1 for ACT. Of course, the fact that we can't identify those who took both tests means that among ACT takers, those from high-income families may be more likely also to take the SAT.

REFERENCES

Bowen, William G., Martin Kurzweil, and Eugene M. Tobin, Equity and Excellence in American Higher Education, University of Virginia Press, 2005.

Bowen, William G. The Thomas Jefferson Foundation Distinguished Lecture Series, University of Virginia, April 2004, Lecture I "In Pursuit of Excellence," Lecture II "The Quest for Equity: "Class" (socio-economic Status) in American Higher Education." Lecture III "Stand and Prosper! Race and American Higher Education."

Cabrera, Alberto F., and Steven La Nasa, "Understanding the college choice of disadvantaged students," New Directions for Institutional Research, 2000, 107, 5-22, 2000

Card, David, and Abigail Payne, "School finance reform, the distribution of school spending, and the distribution of student test scores," Journal of Public Economics, 83, 1, 49-82, January 2002.

Carnevale, Anthony P., and Stephen J. Rose, 2004. "Socioeconomic Status, Race/Ethnicity, and Selective College Admission." In America's Untapped Resource: Low-Income Students in Higher Education, Ed. Richard D. Kahlenberg, New York: The Century Foundation.

Choy, Susan P. (2002). "Students Whose Parents Did Not Go to College: Postsecondary Access, Persistence, and Attainment," from *Condition of Education 2001,* NCES.

College Board, Trends in College Prices (2004).

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

Frank, R. H. (2001). Higher Education: The Ultimate Winner-Take-All Market? Forum Futures: Exploring the Future of Higher Education, 2000 Papers. M. E. Devlin, and Joel W. Meyerson. San Francisco, CA, Josey-Bass, Inc**:** 3-12.

Griffin, Deborah H., "Measuring Survey Nonresponse by Race and Ethnicity," Census Bureau, 2003

Hasseldenz, Jon S. (1976). "Determining Validity and Identifying Nonresponse Bias in a Survey Requesting Income Data," Research in Higher Education, 5, (2) 179-91,.

Heckman, James J. (1999). "Policies to Foster Human Capital," NBER Working Paper 7288.

Hill, Catharine B., Gordon C. Winston, and Stephanie Boyd, "Affordability: Family Incomes and Net Prices at Highly Selective Private Colleges and Universities," Journal of Human Resources, Fall 2005.

Hoxby, Caroline (1997). "How the Changing Market Structure of U.S. Higher Education Explains College Tuition," NBER Working Paper 6323.

Kane, Thomas J., The Price of Admission: Rethinking how Americans pay for college. Brookings Institution Press, Washington DC and Russell Sage Foundation, New York. 1999.

Kayser, Brian D., and Gene G. Summers, "The Adequacy of Student Reports of Parental SES Characteristics," Sociological Methods and Research, 1, 3, 303-315, 1973.

Moore, Jeffrey C., Linda L. Stinson, and Edward J. Welniak, Jr., "Income Measurement Error in Surveys: A Review." Journal of Official Statistics, Vol.16, No.4, 2000. pp. 331-361

Owings, Jeffrey, Marilyn McMillen, John Burkett and Bruce Daniel, "Making the Cut: Who Meets Highly Selective College Entrance Criteria?" Statistics in Brief, April, 1995, National Center for Educational Statistics, U.S. Department of Education, Office of Educational Research and Improvement, NCES 95-732.

Romano, Richard and Louis Moreno, Response Errors in Reports of Parental Income by Community College Freshmen. Working Paper Series No. 1-94. 1994. Institute for Community College Research, Broome Community College, Binghamton, NY

Ross, Catherine E., and John R. Reynolds, "The Effects of Power, Knowledge, and Trust on Income Disclosure Surveys,' Social Science Quarterly, 77, 4, 1996.

Schmidt, Peter, "Noted Higher-Education Researcher Urges Admissions Preferences for the Poor," The Chronicle of Higher Education, April 16, 2004.

Smith, Kris M., Claudia W. McCann, The Validity of Students' Self-Reported Family Incomes. Association of Institutional Researchers, 1998 Annual Forum Paper.

Summers, Lawrence H. 2004a. "Higher Education and the American Dream." Speech delivered at the 86th Annual Meeting, American Council on Education, Miami, Fla., February 29. Available at http://www.president.harvard.edu/speeches/2004/ace.html.

Terenzini, Patrick T., Alberto F. Cabrera, and Elena Bernal. Swimming Against the Tide: the Poor in American Higher Education. College Board Research Report No. 2001-01, 2001.

Vigdor, Jacob L. and Charles T. Clotfelter, "Retaking the SAT", The Journal of Human Resources, XXXVIII:1, 2003.