IS CAMPUS RACIAL DIVERSITY CORRELATED WITH EDUCATIONAL BENEFITS?[1]

by

THOMAS E. WOOD
Executive Director
California Association of Scholars

and

MALCOLM J. SHERMAN
Associate Professor of Mathematics and Statistics
State University of New York at Albany

April 4, 2001

NATIONAL ASSOCIATION OF SCHOLARS

221 Witherspoon Street, 2nd Floor
Princeton, NJ 08542
(609) 683-7878; (609) 683-0316 (fax)
nas@nas.org; http://www.nas.org

THIS DOCUMENT IS AVAILABLE ONLINE AT
www.nas.org & www.calscholars.org

Wood: (510) 843-4957; cas@calscholars.org
Sherman: (518) 442-4628; mjs78@math.albany.edu

[1] Part IV of *Race and Higher Education: Why Justice Powell's diversity rationale for racial preferences in higher education must be rejected*

TABLE OF CONTENTS

TESTING THE POWELL RATIONALE EMPIRICALLY

Viewed as an empirical claim, the Powell rationale asserts that, holding all other variables constant, students benefit by attending racially diverse colleges.

The legal case for Powell's diversity rationale requires more than a bare assertion of First Amendment rights. As universities themselves have come to acknowledge, it must at the very least be supported by empirical evidence that racially diverse student bodies confer educational benefits.

Viewed as an empirical claim, the Powell diversity rationale asserts that, holding all other variables constant, students benefit by attending racially diverse colleges. Testing the hypothesis is contingent on the availability of a sufficiently ample and well-designed database. Given such a database, there is a well-established, standard statistical methodology for testing hypotheses like the Powell diversity rationale, called multivariate regression analysis.

There is enough racial diversity across institutions in the ACE-HERI-CIRP database to test whether students benefit from campus racial diversity. The standard methodology for doing this is multivariate regression analysis.

Testing the Powell hypothesis empirically places significant demands on a database, but as it turns out, there is a database that satisfies those conditions. To its credit, the American Council on Education (ACE), America's largest and most influential higher education organization, started the Cooperative Institutional Research Program in 1966 to construct a higher education database that would be able to answer these kinds of questions.

The ACE is in fact more than an organization; it is, rather, an umbrella group or consortium of organizations with an interest and stake in higher education. Its membership consists of virtually all of the important higher education organizations in the country. The ACE headquarters at One Dupont Circle in Washington, D.C. has become a synonym for the higher education establishment in the United States.

The ACE lists the following as some of its major activities:

- Represents higher and adult education before Congress, federal agencies, the Supreme Court, and the federal courts.
- Conducts research and analyzes data on U.S. higher and adult education.
- Helps shape international education policy at the federal level and works with the national campuses and higher education groups to promote international education.
- Provides opportunities for the exchange of mutual concerns among leading corporate and higher education chief executives.

- Advises colleges and universities in such areas as minority and women's issues, management and leadership, and self-regulation.

- Assists adult learners by administering the General Educational Development (GED) tests and by reviewing and making credit recommendations for learning acquired through courses, programs, and training offered by businesses, labor unions, associations, and the military.

- Publishes news and information through a semimonthly newsletter, Higher Education & National Affairs; a triannual magazine, The Presidency; professional books and guides released through the ACE/Oryx Series on Higher Education; and numerous reports and periodicals.[2]

The American Council on Education (ACE) is much more than an organization; it is, rather, an umbrella group or consortium of organizations with an interest and stake in higher education. The ACE headquarters at One Dupont Circle in Washington, D.C. has become a synonym for the higher education establishment in the United States.

The activity mentioned in the first bullet item -- "representing" or lobbying for higher and adult education before Congress, federal agencies, the Supreme Court, and the federal courts -- is the special responsibility of the Higher Education Secretariat of ACE. The Secretariat was formed on October 5, 1962 to serve as a forum for the chief executive officers in the higher education association community. At present, the Secretariat is composed of 47 national higher education associations, representing the different sectors and functions in postsecondary institutions. The American Council on Education is the coordinating and convening body for the Secretariat. ACE provides a forum for the Secretariat's discussion on education issues of national and local importance. The Secretariat is the chief lobbying arm of the higher education establishment at both the federal and state levels.

The Cooperative Institutional Research Program (CIRP) was founded by ACE in 1966 under Alexander W. Astin, its director of research at the time. Here is how the abstract of the first ACE Research Report on CIRP described the goals of the program:

The Cooperative Institutional Research Program (CIRP) was founded by ACE in 1966 under Alexander W. Astin, its director of research at the time.

> In order to assess the impact of different college environments on student development and to provide a source of current, readily available information about college students, the American Council on Education (ACE) has undertaken a large-scale program of longitudinal research on the higher educational system. The program will be based primarily on a comprehensive file of information from a representative sample of higher education institutions which will be updated annually. ACE's research data file is designed to incorporate the best features of a data base (descriptive information that will

[2] http://www.acenet.edu/About/Membership/how.html. See also American Council on Education 1953.

> eventually become outmoded) and a data registry (information that is stored for future use). It will include longitudinal records on students, with special emphasis on student development, and 4 categories of institutional data: finances and financial policies, curriculum, administrative policies and practices, and faculty. The research data file is designed to serve 3 basic functions: research, information, and training, and will possibly be used by other educational organizations and individuals as a research tool.

When the Bakke case was decided in 1978, CIRP had been in operation for twelve years. In 1977, one year before the Bakke decision, Alexander Astin published the first comprehensive report of the CIRP findings (Astin 1977). At that point, however, CIRP was not in a position to test the Powell diversity hypothesis using multivariate regression analysis, because at that point in time information about the racial diversity of the student bodies had not been incorporated into the database.

This changed in the mid- to late-1980s, when Astin began merging this kind of data, which is available from the U.S. Department of Education's Integrated Postsecondary Data System (IPEDS), along with much other useful information, into the CIRP database.

What Matters in College?: Four Critical Years Revisited (Astin 1993c) was the first, and remains the only, detailed report of the findings from this new, more comprehensive database. Here is how Astin describes the ACE/HERI/CIRP database in his 1993 work:

> CIRP was initiated at the American Council on Education (ACE) in 1966; since 1973 it has been conducted by the Higher Education Research Institute at the University of California, Los Angeles, with continuing sponsorship by ACE. It is now the largest ongoing study of the American higher education system, with longitudinal data covering some 500,000 students and a national sample of more than 1,300 institutions of all types. These data cover a wide range of cognitive and affective student outcomes, affording the opportunity to examine how the college experience affects more than eighty different measures of attitudes, values, behavior, learning, achievement, career development, and satisfaction. The size and scope of CIRP make it possible to employ highly sophisticated multivariate controls over a large number of potentially biasing variables– in particular, the characteristics of the entering students that might

Since 1973 CIRP has been conducted by the Higher Education Research Institute (HERI) at UCLA, with continuing sponsorship by ACE. It is now the largest ongoing study of the American higher education system, with longitudinal data covering some 500,000 students and a national sample of more than 1,300 institutions of all types.

predispose them to pick particular types of colleges or programs.[3]

"The Cooperative Institutional Research Program (CIRP), which has been in progress for nearly a quarter of a century and which now includes input data on nearly 8,000,000 students and 1,300 institutions, was initiated in 1966 specifically to collect input data that would make it possible to apply the [input-environment-output] model to a national study of student outcomes in American higher education. …. it is an omnibus instrument that includes demographic and other background data as well as pretests and self-predictions … on a wide variety of college outcomes …."[4]

THE ACE-HERI-CIRP 1985-89 LONGITUDINAL UNDERGRADUATE STUDY

Astin's *What Matters in College?* reports on the findings of the ACE-HERI-CIRP 1985-89 longitudinal undergraduate study. This study used institutional data from HERI's own Registrars Survey and from the U.S. Department of Education's Integrated Postsecondary Data System (IPEDS). The latter database includes enrollment figures by institution for African American, Asian-American, Latino, and white undergraduate students. It also includes data about institutional finances, degrees earned, faculty salaries, whether the institution is a public or private one, and institutional type (denominational, research university etc.).

"The size and scope of CIRP make it possible to employ highly sophisticated multivariate controls over a large number of potentially biasing variables." --Astin 1993c

The Educational Testing Service (ETS) provided student scores on the SAT, GRE, and NTE (National Teacher Examination). The American College Testing Program, the Association of Medical Colleges, and the Law School Admissions Council provided student test scores on the ACT, MCAT, and LSAT.

Finally, HERI used data provided by the 1989 HERI faculty survey and HERI student surveys. The HERI faculty survey surveyed teaching faculty at 217 of the 309 four-year institutions that were involved in the 1985-89 student survey.[5] The student surveys consisted of the HERI Student Information Form (SIF) for entering students and the 1989 Follow-Up Survey (FUS). The FUS was sent to the 309 SIF institutions in 1989-90. The Exxon Education Foundation and the National Science Foundation (NSF) provided HERI with grants for a weighted follow-up

[3] Astin 1993c: 4.
[4] Astin 1993a: 64.
[5] According to Astin, the HERI sample of faculty respondents is comparable with the population of the faculty survey conducted by the National Center for Education Statistics (1990).

sample from 159 of these institutions. The sample was divided into 26 stratification cells for the purposes of data analysis.

There were 309 four-year institutions and 24,847 students in the database for the 1985-89 longitudinal undergraduate study. The database included 131 student input variables and 135 institutional environmental variables, for a total of 266 control variables. The database included a total of 192 environmental variables–the 135 institutional environmental variables (called "bridge" variables) and 57 intermediate outcome or student involvement variables.[6, 7]

Eighty-two cognitive and noncognitive student outcome variables were used in the study. The study had pretests for 44, or roughly half, of these outcome variables.[8]

Controlling for all possibly relevant control variables is critical to multivariate statistical analysis, as Astin points out.[9] Fortunately, one of the great strengths of the ACE-HERI-CIRP database is the size of the samples and the large number of control variables used. There is no systematic, detailed presentation of all the CIRP variables in Astin 1993c (nor in any of the other HERI-CIRP literature we have seen). However, a rough idea of the student input variables and environmental variables (institutional and student intermediate outcome variables) can be gleaned from the exposition of the findings in Astin 1993c.

The 131 student input variables include 44 pretest measures (from the SIF or Student Information Form), 26 self-predictions, and 61 other input characteristics, including the types of courses taken in high school, preliminary choice of career, the importance given to eleven reasons for going to college, religious preference, parental occupation, parental income, parental education, student race or ethnicity, age, gender, marital status, and citizenship.[10]

The 135 environmental measures include 16 measures of institutional characteristics, 15 measures of curricular requirements, 35 measures of the student's peer environment, 34 measures of the faculty

"Three percentage measures are included in the regressions to assess possible effects of the racial composition of the peer group: African-Americans, Asians, and Latinos. With few exceptions, outcomes are generally not affected by these peer measures, and in all but one case the effects are very weak and indirect."
-- Astin 1993c

[6] There is a discussion of the "bridge" variables at Astin 1993c: 32-70.
[7] There is a brief discussion of the environmental variables, including the "intermediate student outcome" or "student involvement" variables, in Astin 1993c: 365; and a more extensive discussion in Astin's work on statistical methodology (Astin 1993a: 303-308).
[8] Astin 1993c: 14.
[9] (1993c: xxiii-xxiv).
[10] Astin 1993c: 15.

environment, and 35 measures of the freshman's place of residence, financial aid, and choice of major.[11]

THE CRUCIAL TEST

Since the racial diversity of the student bodies in the CIRP database varies from institution to institution, it is possible to use the database to test the Powell rationale empirically using multivariate regression analysis. If the regression coefficient for racial diversity is statistically significant in a model that controls for all explanatory variables, one can say that racial diversity is associated with certain outcomes and is plausibly a cause of these outcomes. If an explanatory variable loses significance when other variables are added to the model, this explanatory variable is said to be an indirect cause of the outcome.

If campus racial diversity enhanced educational quality, one would expect it to have a positive, measurable effect on at least **some** *outcome variables. The HERI-CIRP study found, however, that the racial composition of the peer group had only two direct effects out of a total of 82 outcome variables, both of which were negative.*

Astin reported the findings for the crucial regressions on p. 362 of Astin 1993c. Here is his own description of the findings:

> Three percentage measures are included in the regressions to assess possible effects of the racial composition of the peer group: African-Americans, Asian-Americans, and Latinos. With few exceptions, outcomes are generally not affected by these peer measures, and in all but one case the effects are very weak and indirect. Perhaps the most interesting finding is the negative effect of the percentage of Latino students on attainment of the bachelor's degree. This finding is reminiscent of earlier research … indicating that Chicanos, in particular, are relatively likely to drop out of high school and college, even after controlling for their academic preparation and other background factors. One possibility is that this measure, the percentage of Latino students in the student body, may well be a crude proxy … for the overall dropout rate of the institution.
>
> The only other direct effect is the negative effect of the percentage of Asian-American students on the perception of a Student Oriented Faculty (Beta = -.21). Otherwise, none of these three measures produces any direct effects, and practically all of the indirect effects are very weak.

The California Association of Scholars (CAS) drew attention to this passage in 1993, when it joined a number of member institutions of the Western Association of Schools and Colleges (WASC), including Stanford University, Cal Tech, and a number of denominational colleges, in challenging WASC's "Statement on Diversity." The WASC Statement included the claim that racial diversity is an important

[11] Astin 1993c: 76-77, 81.

element of educational quality and excellence. When challenged on this point, the WASC Senior Commission invoked Astin's work. The CAS was apparently the first to note that the crucial test, which was reported in Astin 1993c: 362, actually disconfirmed the claim that campus racial diversity is correlated with educational excellence.

Since the was apparently the first to draw attention to the importance of the CIRP database to the current national debate over racial preferences in university admissions, we will take the liberty of quoting the relevant passage in full from the CAS' critique of WASC's "Statement on Diversity":

> ["What Matters In College?"] … does not support the claim that ethnic and racial diversity in the peer group enhances educational quality. If this were true, then one would expect that having diverse student bodies would have a positive, measurable effect on at least *some* outcome variables. The HERI-CIRP study found, however, that the racial composition of the peer group had only two direct effects out of a total of 82 outcome variables that were studied, and both of these were negative (ibid., p. 362). …
>
> We are struck by the fact that the study found no correlation between racial diversity in the student body on any of 82 cognitive and non-cognitive outcome variables, and that it actually found, in effect, a *negative* correlation between gender diversity in the peer group and some of the outcome measures. These findings are directly relevant to the most controversial aspect of the current debate over multiculturalism and diversity within the academy– i.e., the insistence on the part of many of the proponents of "diversity" that the university apply race and sex-based preferences in student admissions and promotions, rather than equal opportunity policies requiring that students be admitted and faculty be hired and promoted without regard to the criteria of race, sex and ethnicity. In order to justify a diversity standard for accreditation which would justify the use of such preferences, diversity proponents are obliged to provide empirical evidence showing that diversity is a prerequisite of educational excellence. In our opinion, the HERI-CIRP does not justify this contention. In a number of crucial areas, the study fails to provide the expected evidence that diversity and excellence are positively correlated, and in some areas it actually provides evidence that they may be negatively correlated.[12]

"As it turns out, none of these three measures--considered independently--produced many direct effects on student outcomes. The CAS has seized on these negative findings as a means of refuting the claim … that having a racially diverse student body "enriches" the student's educational experience."
-- Astin 1993c

[12] California Association of Scholars 1993: 30

MITCHELL J. CHANG'S 1996 DOCTORAL DISSERTATION

The CAS' critique of WASC's "Statement on Diversity" created a stir inside WASC, but did not elicit a response from Astin himself (at least not publicly). We were, therefore, surprised to learn this year from a National Association of Scholars member who belongs to another state affiliate that Astin had responded seven years previously to our criticisms.

It is astonishing that Astin has characterized this finding as a "minor" one. Since this is the crucial test for the diversity rationale in multivariate regression analysis, his statement raises the question whether Astin has confidence in his own methodology and database.

Our 1993 critique of the WASC statement referred to the 1991 ACE/Macmillan edition of *What Matters in College?*. In 1993, but apparently after the CAS' critique had been written and distributed, Jossey Bass published the paperback edition of *What Matters in College?* On reading the introduction to the paperback edition of Astin's book recently, we were startled to find the following passage:

> Since the book was first released I have also had two rather sobering experiences concerning the use of these research findings by politicians and by our judicial system. The first of these concerns the current debate over "diversity" and "multiculturalism" on the campus. Within the academy, these attacks have been led by a group that calls itself the National Association of Scholars and, in my own state, by an affiliate called the California Association of Scholars (CAS). While the research findings concerning the effects of multiculturalism are clear--students benefit in a variety of ways when their campus emphasizes multiculturalism in its curriculum and cocurriculum–the CAS has chosen to ignore (or dismiss) these findings in its public pronouncements and to focus instead *on a minor finding concerning racial enrollments* [emphasis ours]. … When we originally devised the 135 environmental measures for this study, we included three simple measures of racial enrollments: the percentages of African Americans, Asian-Americans, and Latinos in the student body. As it turns out, none of these three measures--considered independently--produced many direct effects on student outcomes. The CAS has seized on these negative findings as a means of refuting the claim--typically advanced by proponents of affirmative action--that having a racially diverse student body "enriches" the student's educational experience. The federal courts in the Hopwood case recently reached the same negative conclusion as the one propounded by the CAS in their decision to outlaw racial consideration in admission to the University of Texas Law School.
>
> That these measures do not necessarily reflect "diversity" is easily illustrated by considering the percentage of

> African-American students in the student body. Using this measure, we would have to conclude that the historically black colleges are the "most diverse" institutions, when in fact their student bodies are among the *least* diverse; that is, more than 90 percent of their students are the same racial group! Under these circumstances, it is clearly not reasonable to claim that this study proves that diversity has no consequences for student development.
>
> To explore the issue more directly, Mitchell Chang from our institute recently developed a comprehensive measure of student body diversity as part of his doctoral dissertation. … Under Chang's definition, the most diverse student body would be one with equal representation of students from different racial groups, while the least diverse student body would comprise mostly students from one group. Chang found that white students who attend institutions with diverse student bodies, compared with those who attend institutions enrolling mostly white students, are more likely to discuss racial issues and to socialize with nonwhite students. Since these latter two student experiences are, in turn, associated with a number of positive educational outcomes … it seems clear that diversity can indeed have beneficial effects on student development. It will be interesting to see how the CAS deals with Chang's findings once they are published.[13]

Astin has challenged us to respond to Chang's findings. We are happy to do so here.

It is astonishing that Astin has characterized the results of the regression tests on the racial diversity variable in his study as "minor." Since this is the crucial test in multivariate regression analysis for the diversity rationale, his statement makes one wonder whether Astin himself has confidence in his own methodology and database.

Astin's assertion that Chang found something important that his analysis had simply overlooked is based on two claims. The first is that Chang constructed more precise measures of campus racial diversity that led him to different conclusions. The second is that Chang found that campus racial diversity is correlated with two variables that are correlated in turn with student outcomes.

The first claim is bizarre. Chang 1996 does in fact elaborate three different measures of campus diversity, but the measures do not appear to have led to different findings about the impact of campus racial diversity from those described in Astin 1993: 362. Indeed, we are unable to find any passage in Chang 1996 where he claims otherwise.

[13] Astin 1993c: xvi-xvii.

The second claim does nothing to blunt the force of the fundamental finding of Astin 1993: 362. We note, first of all, that in his introduction to the paperback edition of *What Matters in College?*, Astin cites the finding that white students are more likely to discuss racial issues and to socialize with nonwhite students. But Chang also found, though Astin fails to report, that black students are *less* likely to discuss racial issues or socialize with students of another race on campuses that are more racially diverse.

But this is a relatively minor point. The more important point is that in Astin 1993, these two measures are considered, as surely they must be, as intermediate outcome variables rather than final outcome variables. In short, Chang's new finding comes from treating two of the variables that Astin considered (rightly) as intermediate outcome variables as final outcome variables. So far as the diversity rationale is concerned, the correlation between the SOCIALIZATION and DISCUSSION variables in the CIRP database and final student outcomes are of interest only to the extent that there is a positive synergy or interaction (called an "interaction effect" by statisticians) between the racial diversity of the student body, the SOCIALIZATION and DISCUSSION variables, and final outcomes.[14] But neither Chang nor Astin reports these interaction effects, though as we show below on the basis of indirect evidence, these interaction effects are likely to be very weak.

Chang's Diversity Measures: Much Ado About Nothing

Although we are unable to find any passages in Chang 1996 where he claims that the use of his diversity measures leads to results different from Astin's finding that the racial composition of the peer group had either weak or non-existent effects on student outcomes, Chang does try to make it sound as if the "racial homogeneity" of most of the institutions in his sample makes statistical analysis difficult – a claim that was repeated, as we have seen, by Astin in his introduction to the 1993 paperback edition of *What Matters in College?* Here is what Chang says:

[14] What we have called the SOCIALIZATION and DISCUSSION variables come from the following two items from the 1989 Follow-Up Survey (Wingard et al. 206):

"For the activities listed below, please indicate how often – Frequently, Occasionally, or Not at all – you engaged in each during the past year:

"Discussed racial/ethnic issues"
"Socialized with someone of another racial/ethnic group"

> "[Table 4.2] clearly indicates the lack of racial diversity in the sample, and by extension in our institutions of higher education. … Approximately twenty percent of the institutions in the sample have student populations that are almost entirely composed of one racial group. For half of the institutions, over ninety percent of their student body are of the same racial group. Conversely, only about twelve percent of the sample have less than seventy-five percent of one racial group in their student body. Moreover, of all the institutions, only two percent (8 institutions) have fewer than fifty-five percent of the same racial group. In short, the sample of institutions–much like the U.S. population in general–is skewed toward the side of racial homogeneity. Only a very small percentage of institutions even remotely resemble a diverse campus as defined by this study–namely an equal opportunity for cross-racial interaction for all students. Clearly, such opportunities are limited at nearly ninety percent of the institutions in the sample.[15]

Note that in order for a college to be diverse by Chang's measure, it must give every student, regardless of race, "an equal opportunity to establish cross racial relationships." This means that if there are 5 different races attending a school, in order to get a perfect score on Chang's measure each race would have to comprise 20 percent of the school's students. This would be the mix in Chang's multiracial, multicultural utopia, but it is far from the reality on American campuses. Nor do the demographics of present-day universities pose any problems for statistical analysis.

First of all, the schools in Chang's sample are more white than the general population, but not to the extent that Chang's description would have us believe, as can be seen by comparing his numbers with the data on the 18-24 year old population of the US in 1990.[16] In Chang's sample the median institution had about 90 percent white students. The 18 to 24 year old U.S. population as a whole in 1990 was 70.5 percent white. Since Chang excluded historically black colleges from his sample, this is not too surprising.[17] From Chang's table it appears that 11.6 percent of his sample has 75% or less white students. Another 18.6 percent had between 75% and 85% white students. So, about 30 percent of his sample is about as close to his ideal of "diversity" as we can reasonably expect given the demographics and the existence of historically black colleges (and maybe closer). Another

---

[15] Chang 1996: 87-89

[16] The comparisons are between Changs' numbers in Table 4.2 of Chang 1996: 88 and data provided in the *Historical Statistics of the United States*..

[17] Note that because there are historically black colleges, other colleges will have a lower percentage of black students than they would otherwise have had.

20 percent of the schools have between 85 percent and 90 percent white students – not all that bad, considering the demographic constraints they face. Most importantly, there are plenty of schools, about 30 percent of the sample, that are as diverse as we can reasonably expect a school to be in this society, and they provide more than enough observations to contrast with the least diverse schools to detect meaningful structural diversity effects.

Consequently, the impression that Chang tries to give, that variation in diversity is too small to detect meaningful effects, is quite false - especially given the huge sample size he is using and the ability it confers to detect even very small differences.

This basic point remains valid whether one uses Astin's tripartite measure of the racial composition of the peer group or any of Chang's three measures of racial diversity. Indeed, Chang himself provides the data that clearly shows how close his three different diversity measures are to each other. The correlations between Diversity Range and Diversity Variability were so close that Chang reports findings only for the latter measure. Although he reports findings separately for the Diversity Variability and Diversity Heterogeneity variables, the correlations between these two measures are so close as to be virtually identical, as the following sample of findings shows:

> Institutional type (4 year colleges) is negatively correlated with Diversity Variability (r=-.17; p<.001), and Diversity Heterogeneity (r=-.21; p<.001).[18] Size is positively correlated with Diversity Variability (r=.12; p<.05) and Diversity Heterogeneity (r=.14; p<.01).[19] Selectivity is not significantly correlated with either Diversity Variability (r=.01; p>.05) or Diversity Heterogeneity (r=.06; p>.05).[20] Institutional control is not significantly correlated with either Diversity Variability (private; r=.00; p>.05) or Diversity Heterogeneity (private; r=-.01; p>.05).[21]

Chang's Findings Are Mixed

Chang's doctoral thesis at HERI/UCLA (Chang 1996) sets out six research hypotheses. The principal one for our purposes is Hypothesis 1. Here, in Chang's words, are the results of testing this hypothesis against the database he used (i.e., a subset of the very same database used in Astin 1993c):

Chang's 1999 paper is mainly of interest for reporting how small the correlations are in the CIRP database between the DISCUSSION AND SOCIALIZATION variables and the final outcome variables he considers. Chang found that racial diversity accounts for only 1.1% of the total variance in students' proclivity to socialize with someone of a different race, and only .2% of the total variance in students' proclivity to discuss racial issues.

[18] Chang 1996:93.
[19] Ibid. 95.
[20] Ibid : 97.
[21] Ibid. 1996:

[Hypothesis 1] focuses on the principal goal of the study, that is, to test the educational efficacy of student diversity. Thirteen multivariate regression analyses were conducted to test the educational effects of racial diversity. The hypothesis is only partially supported by the findings. After controlling for student background characteristics, college environmental factors, and college experiences, we find that racial diversity has a positive impact on the white student's inclination to both socialize with someone of a different racial group and to discuss racial issues. This implies that a racially diverse student body is a direct causal factor in how frequently white students socialize cross-racially and discuss racial issues (There was no parallel positive effect observed among students of color, and even the suggestion of a weak negative effect.)

In contrast, some of the findings fail to support the hypothesis, and at times even contradict it. Racial diversity has (a) a negative direct impact on overall satisfaction with college among students of color; (b) a marginal, indirect negative impact on retention among all students; and (c) no effect on intellectual self-concept, social self-concept, or college GPA….

To begin, it is not surprising to find that, the more students of color there are in the peer environment, the more likely white students will be to socialize across racial groups. This would be expected since a diverse student population … creates more opportunities for white students–the majority on almost all campuses–to interact with students of color. What is interesting, however, is that racial diversity has differential effects on students of color and on white students. Since the nonsignificant results for students of color were actually of opposite sign to that found for white students, it seems clear that white students, compared to students of color, are more likely to be positively influenced by a diverse student population.

This phenomenon can be explained in part by changes in the statistical probability for cross-racial opportunities. In a more racially diverse campus there will (by definition) be more students of color, and their numbers will increase opportunities for these students to interact with students of their own racial group. By contrast, increased racial diversity increases the opportunities for cross-racial interaction among white students. Ironically, however, as racial diversity increases and white students subsequently have more opportunities to develop cross-racial friendships, students of color, if anything, become more inclined to socialize with some of their own race. …

In order to assess the likelihood that future research might demonstrate direct effects between racial diversity and final outcome variables, one needs to know what the interaction effects are between input variables, intermediate variables, and outcome variables. Quite unaccountably, HERI has not reported these effects. The likelihood that these crucial numbers are very low probably explains why they are not reported.

> The results also show that diversity has a negative effect on overall satisfaction among students of color and are suggestive of a weak, indirect negative effect on college retention for all students. Moreover, diversity does not affect students' college GPA, intellectual self-concept, or social self-concept. It appears then that the effects of racial diversity are inconsistent, if not paradoxical. On the one hand, racial diversity has a negative impact on college satisfaction and possibly also on retention. On the other hand, it is positively associated with variables that themselves have positive effects on these very same outcomes."[22]

Chang's Findings are Extraordinarily Weak

In discussing Chang's findings, it is necessary to refer to two different writings. One is Chang's doctoral dissertation of 1996. The other is an article by Chang that was published three years later.[23] Chang 1996 and Chang 1999 use different databases. Each database, however, is a subset of the full CIRP database of the 1985-89 longitudinal survey that was used by Astin in Astin 1993c. The database used in Chang 1999 is the smaller of the two subsets, as shown in the following:

(1) The database for Astin 1993c included 309 four-year institutions; 24,847 students; 131 student input variables; 135 environmental "bridge" variables; 57 student involvement variables; and 82 final outcome variables.

(2) The database for Chang 1996 included 371 four-year institutions; 11,600 students, 26 student input variables; 24 environmental "bridge" variables; 13 student involvement variables; and 7 final outcome variables.

(3) The database for Chang 1999 included 371 four-year instituitons; 11,688 students, 16 student input variables; 14 environmental "bridge" variables; 14 student involvement variables; and 6 final outcome variables.[24]

Chang 1999 essentially adds nothing to Chang 1996. The 1999 paper is mainly of interest because it reports how small the correlations are between the DISCUSSION AND SOCIALIZATION variables and the final outcome variables:

> "When entered in the equation after controlling for precollege variables, racial diversity [of the institution] was associated with

[22] Chang 1996: 149-155.
[23] Chang 1999.
[24] See Appendix VII for a tabular comparison.

> an R^2 increase of 1.1% [i.e., the racial diversity of the institution accounts for 1.1% of the total variance in students' proclivity to socialize with someone of a different race]. Although the total variance accounted for by this equation is not remarkable … racial diversity continued to significantly affect student opportunities to socialize across race even after controlling for relevant student background, college environments, and college experiences."[25]
>
> "Similarly, racial diversity can be said to have a direct positive effect on discussed racial issues. After controlling for precollege characteristics, racial diversity accounted for an additional .2% R^2 increase [i.e., the racial diversity of the institution accounts for .2% of the total variance in students' proclivity to discuss racial issues].[26]

Obviously, these correlations are minuscule.

Interaction Effects, Intermediate Outcome Variables, and Final Outcome Variables

As we have previously noted, Chang (1996 and 1999) classifies the DISCUSSION and SOCIALIZATION variables as final outcome variables, whereas Astin classifies them as intermediate, student involvement variables. Astin's classification is clearly the preferable one.

Chang found that on campuses with greater racial diversity, white students are more likely to socialize with nonwhite students and to discuss non-racial issues. Note, however, that the increased discussion of racial issues reported by white students need not necessarily have been with nonwhite students. Furthermore, increased socialization by white students with nonwhite students is exactly what one would expect simply on the basis of Brownian motion on campuses that have more racial diversity, unless those campuses are racially segregated or balkanized. As Chang himself puts it:

> " …. it is not surprising to find that, the more students of color there are in the peer environment, the more likely white students will be to socialize across racial groups. The would be expected since a diverse student population … creates more opportunities for white students - the majority on almost all campuses - to interact with students of color."[27]

---

[25] Chang 1999: 388.
[26] Chang 1999: 388-89.
[27] Chang 1996: 150.

So far as testing the Powell rationale is concerned, mere discussion of racial issues (across or within racial lines) and socialization across racial lines cannot be regarded as ends in themselves. At most, they could be regarded as desirable outcomes only to the extent that they have (as one might hope) a positive impact on other outcomes, which is presumably why Astin in Astin 1993c considered them as intermediate, student involvement variables.

As we have seen, in his introduction to the paperback edition of Astin 1993c, Astin said: "Since these latter two student experiences are, *in turn* [emphasis ours], associated with a number of positive educational outcomes … it seems clear that diversity can indeed have beneficial effects on student development." Note that this is something that it would not be appropriate to say about real final outcome variables, like satisfaction with college, GRE scores, or propensity to go on to graduate work. Take satisfaction with college, as an example. While it might make sense to test for correlations between satisfaction and say, propensity to go on to post-graduate work (i.e., to test the hypothesis that all other things being equal, students who are satisfied with their overall college experience tend to stay longer in higher education than those less satisfied), a measure of satisfaction in college is clearly "final" in a way that socialization across racial lines or discussing racial issues (across racial lines or not) is not. In undesirable campus climates, the impact of these two CIRP variables could actually be negative, and universities should take steps to make sure that they are positive. But all universities will, for example, as a matter of course want their students to be satisfied with their college experience.

One must keep in mind that as far as the Powell rationale is concerned, the question is not whether DISCUSSION and SOCIALIZATION have effects on real, final outcomes, but whether RACIAL DIVERSITY does. That is, what one really wants to know from the CIRP studies is how much of a difference, if any, the racial diversity of a student body has on real final educational outcome variables, given that the racial diversity of a student body has a statistically significant impact on DISCUSSION and SOCIALIZATION. This would give some indication of the likelihood that universities might at some point be able to improve the quality and frequency of discussions of race and cross-racial socialization on campuses enough to eventually be able to to demonstrate that campus racial diversity has an educationally significant impact on final student outcomes.

While Chang gives in detail the step-by-step beta coefficients for the DISCUSSION and SOCIALIZATION variables, neither he nor Astin provides the really crucial data, at least so far as the Powell diversity

The "in turn" language used by the HERI-CIRP researchers shows clearly that the DISCUSSION and SOCIALIZATION variables are really intermediate variables rather than final outcome variables. That is exactly how they were treated in Astin 1993c. Chang, however, prefers to treat them as final outcome variables, probably in an effort to find ***some*** *correlation between final outcomes and racial diversity.*

hypothesis is concerned, which is the way that these variables interact with racial diversity and final educational outcomes.[28]

An interaction effect is essentially a measure of the positive or negative synergies between (in this case) the following potentially interacting variables: the racial diversity of the student body, the DISCUSSION and SOCIALIZATION variables, and the final student outcome variables. The ACE-HERI-CIRP literature frequently reports that the researchers examine interaction effects, as is quite typical in multivariate analysis research generally. Furthermore, the interaction effects are precisely the phenomenon to which Chang is drawing our attention when he refers to the "potential" that racial diversity has to produce beneficial final student outcomes. In any case, the interaction effects are certain to be very weak, since correlations that would be used in the calculations are very small, and because one gets the interaction effects essentially by multiplying probabilities (i.e., fractions). The likelihood that the crucial numbers would be very low probably explains why they are not reported.

RACIAL DIVERSITY AND VIEWPOINT DIVERSITY

According to Justice Powell, preferential admissions policies like the plan that Harvard was following in 1978 can meet the compelling state interest test because campus racial diversity is correlated with viewpoint diversity and with beneficial educational outcomes. According to Powell, this provides a compelling justification for such policies under the First Amendment, even though he recognized that these were in obvious "tension" with Fourteenth Amendment principles. In his opinion in Bakke, Powell said:

> It is the business of a university to provide that atmosphere which is most conductive to speculation, experiment and creation. …. 'The Nation's future depends upon leaders trained through wide exposure to that robust exchange of ideas which discovers truth `out of a multitude of tongues, [rather] than through any kind of authoritative selection.' United States v. Associated Press, 52 F. Supp. 362, 372."

[28] Astin discusses interaction effects, and stresses their importance, in Astin 1993a, his systematic work on methodology (1993a: 67, 120-127, 298-300, 311-312). Researchers studying longitudinal multivariate databases examine interaction effects as a matter of course. Since the defense of racial preferences in university admissions has come to rest primarily on what might be called the "in turn" hypothesis, which essentially involves an hypothesis about interaction effects, it is odd that, at least to our knowledge, the ACE–HERI-CIPR literature fails to report the interaction effects that the researchers believe justify (or might someday justify) preferential admissions policies.

> The atmosphere of "speculation, experiment and creation" - so essential to the quality of higher education - is widely believed to be promoted by a diverse student body. … As the Court [438 U.S. 265, 313] noted in Keyishian, it is not too much to say that the "nation's future depends upon leaders trained through wide exposure" to the ideas and mores of students as diverse as this Nation of many peoples.
>
> Thus, in arguing that its universities must be accorded the right to select those students who will contribute the most to the "robust exchange of ideas," petitioner invokes a countervailing constitutional interest, that of the First Amendment.

While this passage is not as clear as one might wish it to be, it is reasonably clear that the following different, though related, claims are packed into it:

(1) Racial diversity is correlated with greater viewpoint diversity on campus.

(2) Viewpoint diversity, with or without a racial dimension, is an intrinsic educational good, and one that has beneficial educational outcomes.

(3) Even beyond this simple correlation, racial diversity adds another *dimension* to viewpoint diversity on a campus. Statistically speaking, this amounts to the claim that racial diversity *interacts* with viewpoint diversity.

(4) Racially dimensioned viewpoint diversity is a compelling justification for seeking racial diversity on a campus, and justifies its pursuit even by means that are in “tension” with the Fourteenth Amendment.

(5) Racially dimensioned viewpoint diversity provides a constitutional justification for preferential admissions policies that could not be provided by correlations between racial diversity and other effects (e.g., increased frequency of interracial socialization, interracial dating, or interracial marriage on campus).

The CIRP database is robust enough to throw light on claims (1)-(3). So far as we know, no research based on the database has yet addressed claims (2) and (3), but Mitchell J. Chang has looked at the first claim in an unpublished paper entitled: “Does increase [sic] racial diversity lead to a more diverse collection of thoughts, ideas, and

opinions on campus?: A study of racial diversity and students' viewpoints."[29]

In the Abstract of the paper, Chang poses the following three questions (all three of which he believes are to be answered in the affirmative):

(1) Do viewpoints that are meaningful to higher education differ by racial groupings?

(2) Do campuses that are more racially diverse have a broader collection of viewpoints held by students?

(3) Do the long-term trends of student's collective viewpoints vary by the level of racial diversity on campus?

The statistical tests that Chang used are ANOVA, Kruskal Wallis, and the Levene test of homogeneity of variances. The tests were applied to a national sample of over 5,000 students from 93 four-year institutions. The students were surveyed in the fall of 1994 and again in 1998.[30]

The analysis in "Diversity & Viewpoints" is based on an examination of the responses to seven questions on the 1994 CIRP entering freshmen survey (SIF) and 1998 follow-up survey (FUS). This selection was based, presumably, on the assumption that the questions might reasonably be expected to show diversity of opinion across racial groups. The questions were as follows:

(1) The Federal government is not doing enough to protect the consumer from faulty goods and services.

(2) The Federal government should raise taxes to reduce the deficit.

(3) There is too much concern in the courts for the rights of criminals.

(4) The death penalty should be abolished.

[29] Henceforth we shall refer to this paper by its running head: "Diversity & Viewpoints." The paper is available on the Web as an Adobe Acrobat PDF document at: http://www.gseis.ucla.edu/faculty/chang/Diversity_Viewpoints.pdf

[30] This is a different sample from the one used in Astin 1993, Chang 1996, and Gurin 1999. Astin 1993, Chang 1996, and Gurin 1999 use the 1985-1989 ACE-CIRP undergraduate study database (or a subset of it); Chang's "Diversity and Viewpoints" uses the 1994-98 ACE-CIRP undergraduate study database.

(5) Employers should be allowed to require drug testing of employees or job applicants.

(6) A national health care plan is needed to cover everybody's medical costs.

(7) Racial discrimination is no longer a major problem in America.

(8) Wealthy people should pay a larger share of taxes than they do now.

(9) Colleges should prohibit racist/sexist speech on campus.

Chang also used a measure of student Political Orientation. This was also taken from the CIRP 1994-98 undergraduate study. The Political Orientation measure used a five-point scale: far left; liberal; middle of the road; conservative; and far right.

Increasing racial diversity is not correlated with increasing viewpoint diversity in any consistent way. Furthermore, what changes there are are extremely small, and show no consistent directionality.

According to Chang, the findings support the conclusion that racial diversity is a "reasonable proxy" for viewpoint diversity in higher education:

> The results of this study correspond well with Justice Powell's intuition that attaining greater diversity broadens the range of viewpoints held by the student body. According to him, these differences in viewpoints are meaningful for higher education because they stimulate a more robust exchange of ideas, which encourages greater speculation, experimenting, and creativity. …. This study helps to clarify how diversity may actually foster those benefits by identifying several attributes that are unique to racially diverse student bodies. … Because racial differences at the group level continue to be meaningful and perhaps even intractable, the opportunity to learn from and with students of different races is educationally compelling and relevant.
>
> The results of this study reveal that accounting for race to admit students achieves more than a student body that 'looks different.' At the point of college entry, there are significant differences in viewpoints between racial groups on a variety of pressing contemporary issues. Although individuals of any given race hold the full range of opinions, as a group, average viewpoints differ from each other. These racial group differences are educationally relevant because they affect the collection of thoughts, ideas, and opinions of the student body. Overall, the variance of students' collective viewpoints tends to be greatest in the most racially diverse campuses, followed by somewhat diverse, then least diverse campuses. These findings suggest

> that more racially diverse campuses have a *much higher likelihood* [emphasis ours] of exposing their students to a broader range of viewpoints on academically relevant issues."[31]

This passage addresses only one of the claims that we believe are packed into the Powell rationale for racial preferences in university admissions -- the claim that racial diversity is correlated with greater viewpoint diversity on campus. Even on this point, however, we must disagree with Chang.

Since this is a matter of judgment, readers will have to decide for themselves whether Chang's findings justify the claim that race serves as a "reasonable proxy" for viewpoint diversity on a campus.[32] Readers should probably begin by looking at Figures 1 and 2 from Chang's paper. Figure 1 plots the responses of students, broken down into the categories white, black, Asian-American, and Latinos, against the Political Orientation variable. Figure 2 plots the responses of students, broken down by the same racial/ethnic categories, against the statement: "A national health care plan is needed to cover everybody's medical costs."

Chang's findings show that trying to enhance viewpoint diversity on a campus by increasing the racial diversity of the campus would be extraordinarily inefficient.

Chang uses standard statistical tests to test for the statistical significance of group variances, and finds that the differences in variances are significant for some (though not all) of these cases. But such tests do not settle the really crucial question: When are the differences in variances *educationally* significant?

There is no well-established standard for answering this kind of question. It is a matter of judgment. We are constrained to say, however, that when we view these graphs, we are struck far more by the overlap and the similarity between the curves than by the differences.

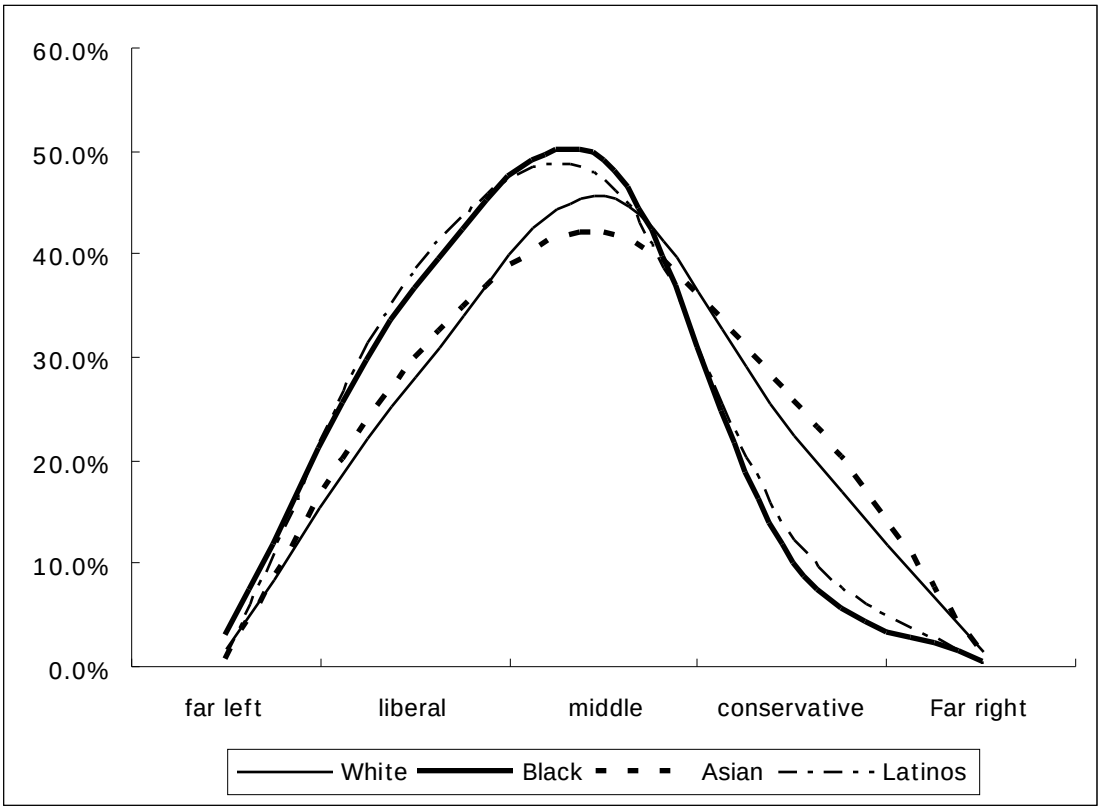
---

[31] "Diversity & Viewpoints," pp. 35.
[32] Ibid., p. 5.

From Chang, "Diversity & Viewpoints," p. 48, http://www.gseis.ucla.edu/faculty/Diversity_Viewpoints.pdf

Figure 1

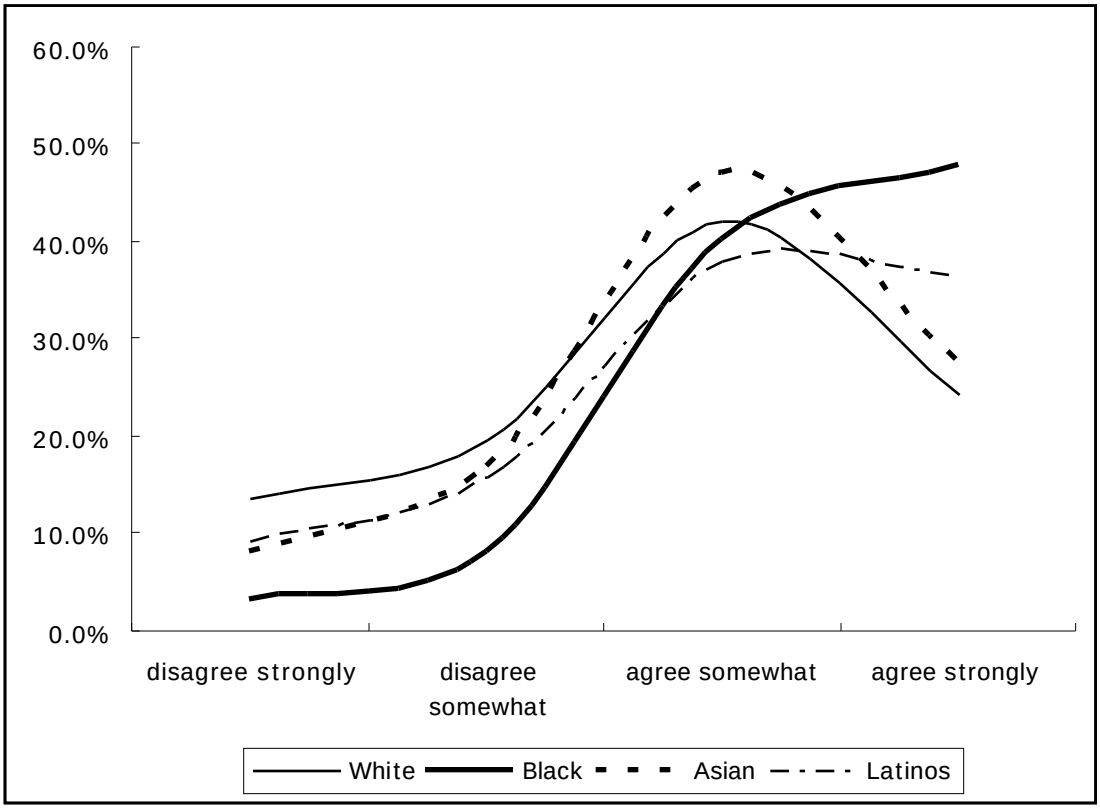
Political Orientation by Race



From Chang, "Diversity & Viewpoints," p. 49, http://www.gseis.ucla.edu/faculty/chang/Diversity_Viewpoints.pdf

Figure 2

Opinion on the Need for a National Health Care Plan by Race



One way to get a better handle on the question is to look at the means and the size of the standard deviations. Table 2 on p. 43 of Chang's paper (which is not given in this report) shows that the mean and standard deviation differences are very small. Furthermore, increasing racial diversity is not correlated with increasing viewpoint diversity in any consistent way. The average change in the variances of viewpoints between 1994 and 1998 by level of racial diversity for the ten items Chang examined are as follows ("Diversity & Viewpoints," Table 6, p. 48):

	Low Diversity	Medium Diversity	High Diversity
(Average of the 10 items)	-0.019	0.021	-0.001

These changes are extremely small. Furthermore, they show no consistent directionality. The average change of variances for "Medium Diversity" institutions is positive (though very low), whereas the average change of the variances for "Low" and "High Diversity" institutions is negative (though also very low).

Chang did not attempt to assess the impact that viewpoint diversity has on final student outcomes, although this would clearly be worth doing.

Another way of looking at the question of educational significance is to ask how efficient it would be to try to enhance viewpoint diversity on a campus by increasing the racial diversity of the campus. It is evident from the graphs and from the tables in Chang's paper that this would be an extraordinarily *inefficient* means of increasing viewpoint diversity.

The simplest and most straightforward way to do this would be to have universities simply question applicants about their viewpoints on various issues, and to consider their responses in their admissions decisions. But there are other possibilities. One is to use socioeconomic status (SES) as the proxy variable for viewpoint diversity (much in the way that some have argued that SES should be used as a substitute for race in admissions policies generally). Since we know from the CIRP database itself that SES is a highly salient variable in higher education, there is good reason to believe that SES would be a much more efficient criterion to use to promote viewpoint diversity than racial diversity. Chang, in fact, mentions this as a possibility that needs to be investigated, though he does not do so in "Diversity and Viewpoints."[33]

So far, we have used Chang's findings to examine only the second of the four questions we believe are packed into the Powell diversity rationale (i.e., whether campus racial diversity is correlated with viewpoint diversity). The other two questions, however, involve

[33] Ibid. p. 19

assessing the impact that viewpoint diversity has on final student outcomes. As Chang acknowledges, the tests and methodology he used in "Diversity and Viewpoints" cannot address these questions:

> This study … is not a college impact study even though longitudinal data were utilized. The purpose of the analyses was not to investigate how increased racial diversity might affect students' educational experiences as other studies have previously done … but to examine whether there are differences between racial groups and between students enrolled in institutions with differing levels of racial diversity. Accordingly, this study used designs with strength in testing between group differences and for some cases, with ability to handle significant differences between group variances. … The designs used in this study, however, do not control for student background characteristics, institutional factors, and student experiences that are typically accounted for in college impact studies. It is unclear if and how these other factors might affect the dependent measures and examining these effects would be an important area for future studies.[34]

There are very good reasons for using the CIRP database to test for college impacts. One is that such tests would provide an additional (albeit indirect) way of testing for the educational significance of the differences in variances themselves. That is, any showing that variances in viewpoint diversity are correlated with final student outcomes in an educationally significant way would be evidence for the *educational* significance of those variances themselves. Another is that impact tests would probably come much closer than the ones Chang has already employed to testing the claims that lie at the heart of Powell's diversity rationale for racial preferences in university admissions.

Using the CIRP database to get closer to the inner core of Powell's diversity rationale would involve doing three things:

> Constructing a measure of viewpoint diversity, using the items provided by the CIRP database.
>
> Finding the correlations between viewpoint diversity and final student outcomes.
>
> Determining the *interactions* between racial diversity of the student body, viewpoint diversity, and final student outcomes.

[34] Ibid., pp. 18-19.

We doubt very much that these tests would support Powell's diversity rationale for racial preferences. For one thing, as even Chang concedes, the overlap in the responses to the ten items is very large. This means that the correlations between viewpoint diversity and final student outcomes are likely to be very small as well. Furthermore, since HERI has found that there are no educationally significant correlations between racial diversity and final student outcomes, it is likely that the interaction effects between racial diversity, viewpoint diversity, and final outcomes will also be negligible or non-existent.

THE EXPERT WITNESS REPORT OF PATRICIA Y. GURIN

Patricia Y. Gurin is Chair of the Department of Psychology and Interim Dean of the College of Literature, Science, & the Arts at the University of Michigan at Ann Arbor. When the University became a defendant in a lawsuit brought by the Center for Individual Rights and Maslon Edelman Borman & Brand on behalf of plaintiffs Jennifer Gratz and Patrick Hamacher, the University commissioned Prof. Gurin to write a report as an expert witness in the case defending the educational benefits of racial diversity on campus. The report she submitted to the court is the "Expert Witness Report of Patricia Y. Gurin" (Gurin 1998). Our references to this document (which we will sometimes refer to as the "Gurin Report") are to the hard copy version submitted to the court. An online version without pagination is available on the Web at http://www.umich.edu/~urel/admissions/legal/expert/gurintoc.html.

Three different studies are incorporated in Gurin's Expert Report: a study based on subset of the database for the HERI-CIRP 1985-89 longitudinal undergraduate survey; a University of Michigan study (the Michigan Student Study, or MSS); and a study based on a course given at the University by the Intergroup Relations, Conflict, and Community Program (IRCCP). In this section of our report, we shall be concerned only with the first study.

Of the studies and reports dealing with the 1985-89 HERI-CIRP longitudinal undergraduate study, the fullest by far is Astin 1993c. That study involves 309 four-year institutions; 24,847 students, 131 student input variables, 135 environmental variables, 57 student involvement variables, and 82 outcome variables. The later studies by Chang (1996 and 1999) involved a subset of this 1985-89 undergraduate database.[35] The database used in Gurin's Expert Report was an even smaller subset than was used in either of the Chang studies. The Gurin study involves 184 four-year institutions, 9,316 students, 6 input

[35] See the composite table in Appendix VIII of this report.

variables, 6 environmental "bridge" variables, 5 student involvement variables, 15 learning outcomes, and 8 democracy outcomes.[36]

On 12 February 1999, Gurin was deposed by David Herr of Maslon Edelman Borman & Brand.[37] We learn from her deposition that HERI was, in effect, a cooperating organization in the production of the Gurin Report. Gurin was not herself familiar with CIRP (Gurin 1999: 10) and therefore relied on Eric Dey and Sylvia Hurtado, two colleagues at Ann Arbor who were familiar with it (Gurin 1999: 8). Both Dey and Hurtado received their doctorates from HERI/UCLA, and in fact were senior researchers there before joining the faculty of the University of Michigan.[38] It was Dey who went to UCLA to get permission to use the CIRP data set (Gurin 1999: 27-28). The computer analyses were also conducted by Dey (as well as by some students). In her deposition, Gurin did not say how the institutional sample was selected from the CIRP database, other than that she asked Dey to "specify our schools by schools that represent the mass number of schools in the United States where structural diversity can be more operative" (Gurin 1999: 41). During the deposition, she acknowledged that she did not know which institutions were included in her study (Gurin 1999: 41), or even whether the University of Michigan itself was (Gurin 1999: 38).

Since her Expert Report was written to defend the University of Michigan's racially preferential admissions policies, the key variable for Gurin's study was what Astin (1993c) calls the "racial composition of the peer group," or what Gurin calls the "structural diversity" of the institution.

An Overview of the Gurin Report

The central problem that Gurin faced in producing her Expert Report is that the national database on which she had to rely actually *disconfirms* the claim that she was asked by the University to defend. How does Gurin deal with this highly embarrassing fact?

The short and definitive answer is that *she tries desperately to change the subject*. Her whole methodology is to treat what Astin calls "diversity activities" (and what she calls "campus experience variables") as the real matter of interest, rather than the racial diversity of the institution (what Gurin calls "structural diversity").

As a way of approaching the only question before the court, this is absurd, and can be explained only as a devious attempt to distract attention from Astin's fundamental finding that there are no educationally significant correlations between racial diversity and student outcomes.

---

[36] These figures are for the four-year study only. The part of the Gurin Report that is based on the CIRP data also includes a nine-year follow-up study, which involved different outcome variables. See the composite table of Appendix VIII of this report.

[37] Gurin 1999.

[38] In recent years, the University of Michigan has placed a major emphasis on "diversity." For an interesting discussion of the University's diversity program, see Lynch 1997: 274-323.

But there are other problems with Gurin's analysis as well:

First, Gurin uses a much smaller subset of the full and comprehensive database that is available for the CIRP 1985-89 longitudinal study, incluiding far fewer students and control variables than the full CIRP study. The inclusion of far fewer control variables is a particularly serious matter, since the validity of regression analysis depends on controlling for all possible sources of bias. Gurin does not justify her omission of two year colleges from the CIRP database. Gurin also drops historically black colleges from the CIRP sample. While "percent minority" has a different meaning at these colleges than at other institutions, the data from historically black colleges is still relevant to some peripheral issues. Instead of dropping these institutions from the sample, Gurin could have analyzed these data separately.

Although Gurin includes far fewer students and institutions than were available to her, the samples she used were still very large. Huge databases of the sort she worked with are capable of showing trivial effects as being statistically significant. That is why Astin himself follows the practice of not reporting on statistically significant findings unless the Beta coefficients–a measure of the predictive power as opposed to the statistical significance of a variable–have an absolute value of .15 or higher. Without any explanation, Gurin departs from this practice, and reports anything with a statistical significance of .05, no matter what the magnitudes–i.e., Beta coefficients–might be. In fact, in order to suggest more statistically significant relationships for black and Hispanic students, she uses an even more liberal test of statistical significance for them ($p<.10$).

Her modus operandi is to then place little black boxes in her tables whenever she finds a purely statistical significance for her relationships. This is clearly intended to give the statistically unwary the impression of a consistent pattern of significant findings. In fact, the findings are trivial. In many cases, the relationships are not even statistically significant, and in the cases where Gurin claims statistical significance, the correlations are entirely trivial. Typically, the measures she studies explain less than 2% of the total variance. These magnitudes are so low that they could not possibly be taken by any reasonable person as a justification for racially discriminatory admissions policies.

But the situation is even worse than this, because even Gurin's claim of statistical significance is suspect, given the type of regressions she employs. She considers four diversity experiences: attending a racial/ethnic workshop, discussing racial issues, socializing with someone of a different race, and having close college friends of a different race (or of the same race, in the case of minority students). Gurin never explains why she uses models that utilize only *one* of

Gurin's term "classroom diversity" is likely to be misleading, since it invites confusion with the very different concept of "racial diversity in the classroom." There is no necessary connection between Gurin's "classroom diversity" variable and racial diversity, since it is entirely possible for a white student, for example, to take an ethnic studies course that has no blacks and Latinos in it at all.

these four variables at a time. This is an anomalous way to model a regression analysis. Using the standard model of regressing on all four variables at the same time would very likely reduce the number of relationships in her study that are even statistically significant, though we are unable to say by how much, since the database she uses is not available to us.

Gurin's claims of statistical significance are further impeached by the fact that her outcome variables are soft student self-assessments, many of which are non-cognitive "democracy outcomes."[39] There may be only a weak relationship between student self-reports for these outcomes and real outcomes, and in any case the relationships are certainly far from perfect. (This is also true for outcomes like self-reported GPAs, but the problem is especially acute for the kinds of non-cognitive outcomes that are at the heart of Gurin's report.) Self-reported variables are not criterion-based variables; one must therefore factor in the weakness of the relationship between the self-reported variable and real outcomes. This consideration is a further impeachment of Gurin's claims even for the purely *statistical* significance for her relationships.[40]

The most important point, however, and one that cannot be emphasized enough, is that the relationships Gurin chose to study are what she calls "diversity experience variables" rather than campus racial diversity itself. As we have already pointed out, this makes absolutely no sense as a way of addressing the question whether campus racial diversity might produce educational benefits (and therefore whether racial preferences in admissions might do so by increasing racial diversity through brute force). After all, the University of Michigan was not sued because it was offering ethnic studies courses or workshops, or because it had prohibited discussion of race on campus or interracial socializing. It was taken to court because its admissions policies employ racial classifications that are designed to increase racial diversity.

---

[39] Many would question whether Gurin's "democracy outcomes" are genuine academic outcomes. Even among the self-reported cognitive outcomes, only one–self-reported GPA–asks the student to report on the *University's*, rather than the student's, assessment of his or her academic skills.

[40] Astin, of course, never claimed to find any educationally significant correlations between racial diversity and student outcomes, so the question doesn't even arise here. Nevertheless, it is worth pointing out that Astin did claim in *What Matters in College?* that there are educationally significant correlations between diversity activities (what Gurin calls campus diversity experiences) and some non-cognitive outcomes of the kind that Gurin calls "democracy outcomes." The point we are making here, therefore, applies with equal force to Astin's claims about the educational benefits of his "diversity activities." In theory, the same weakness of correlation might lead to findings of non-significance when the real correlation is significant, but this is only an abstract possibility; student self-reports are likely to find positive relationships where none exist, rather than the other way around.

This is obvious enough just as a purely intuitive or a priori matter. However, Gurin's methodology is rendered all the more inexplicable because internal evidence from the CIRP database itself shows that diversity activities cannot be treated as even a rough proxy for racial diversity. The reason is, simply, that Astin's study, which uses the full CIRP database, found some educationally significant correlations between diversity activities and some student outcomes, but *none* for racial diversity itself.

Gurin has attempted to deflect this fundamental and dispositive objection to her entire methodology by claiming that it is still useful to consider diversity experiences, because they are correlated with educational benefits, and because the diversity experiences themselves are correlated "in turn" with campus racial diversity.

This response to our objection is utterly futile, for at least two reasons. First, the argument is unsound, because it can be shown mathematically that if variables A and B are positively correlated, and variables B and C are positively correlated, it is possible that A and C are negatively correlated. To be sure, one can deduce that A and C are positively correlated if one knows that the correlations between A and B and between B and C are very high (near 1). But Gurin's correlations are less (usually much less) than 0.25, which isn't nearly large enough to justify the in-turn argument logically or mathematically.

But there is more to this than the purely mathematical argument. The point is further confirmed by internal evidence from the very database that Gurin herself uses. That is because the four campus experience variables that Gurin considers (i.e., the things that Astin typically calls "diversity activities") are *controlled for in the ACE-HERI-CIRP regressions.* After all these variables have been controlled, the regressions fail to find significant correlations between racial diversity and final student outcomes." This means that Gurin cannot argue that racial diversity produces educational benefits *even when it is conjoined with these other factors*. This finding completely devastates Gurin's "in turn" hypothesis.

In order to deflect attention from the central methodological error of her analysis, Gurin has also alluded to what she calls the "remarkable consistency" in the pattern of statistically significant relationships. But this attempt to meet the objection is also futile, for several reasons.

First, Gurin's assertion that statistical significance is a "basic indicator of the strength" of her relationships is a fundamental–and rather elementary–statistical error. With a large enough sample, even a small difference can be statistically significant (i.e., hard to explain by the luck of the draw), but this doesn't necessarily make it important.

The more serious objection, however, is that the pattern of findings to which she refers is totally irrelevant to the matter before the court. That is because the pattern to which she refers is not between student outcomes and campus racial diversity; it is instead between student outcomes and four diversity activities, none of which can be treated as even a rough proxy for racial diversity.

It is, therefore, inexplicable, and indeed rather astonishing, that Gurin should ask us to eyeball a pattern of "consistent findings" about irrelevant variables in order to somehow convince ourselves that racial diversity is correlated with beneficial educational outcomes. After all, the CIRP database that she herself uses permits a direct test of the *relevant* hypothesis. Astin tested it on what is essentially the same database and concluded that one must *reject* the hypothesis that they are so correlated–a fundamental and dispositive finding to which Gurin never refers even once in her entire report.

Gurin is actually at two removes from the policies she claims to be defending. As we have seen, the variables she considers (diversity experiences) are not relevant to the issue before the courts. But even a showing that racial diversity per se *is* correlated with educational benefits would not settle the matter, for it matters *how racial diversity is achieved*. It also has to be demonstrated that racial preferences, which are designed to artificially increase diversity, do not themselves have a *negative* impact on student outcomes.

Although Gurin's diversity experience variables cannot serve as proxies for campus racial diversity, it is often useful to look for proxies for the variables one wants to study. We believe this may be true for racial preferences in university admissions. In order to test the hypothesis that racial preferences have a negative impact on student outcomes, one proxy variable that should probably be examined is the selectivity of the institution, especially as measured by the SAT and other standardized tests. Much of the current controversy over affirmative action in university admissions centers on these tests, since their use has an adverse impact on black and Hispanic applicants. Since the disparate impact is likely to be greater the more selective the institution, any correlations involving Selectivity (a variable that is include in the HERI-CIRP database) is of considerable interest. And in fact Astin (1993c: 51) reports that there is a strong correlation (.39) between Selectivity and Racial Tension. This suggests that it might be useful to examine the possible impacts of racially preferential admissions policies on colleges and universities by using Selectivity as a rough proxy or substitute for the variable Racial Preferences.

This is exactly the kind of finding that researchers should be able to explore with the HERI-CIRP database itself. Unfortunately, the HERI-CIRP database has not been released to any researchers who do not

themselves favor racial preferences. The fact that the HERI-CIRP database has not been made widely available to other researchers is a radical and disturbing departure from normal research practice. It is particularly disturbing now that the CIRP database is at the very center of controversies over racially preferential admissions policies in what is potentially landmark constitutional litigation.

Like the much larger and more comprehensive HERI-CIRP database from which it is drawn, the database that Gurin uses has not been made available to the wider research community either. Her omission is even more indefensible than HERI's. In *What Matters in College?*, Astin was looking at so many ways in which student and institutional characteristics might affect a large number of possible student outcomes that it would have been unreasonable to expect him to provide the data and methodological details that would be expected in a scholarly, peer-reviewed research report on a very specific research question.[41] But the omission in Gurin's case is much more serious, for in her "Expert Witness Report" she is not trying to paint a picture of the college experience generally with broad strokes; she is instead addressing (or more precisely, should have addressed) a very specific and narrow research question. Her research report should therefore have been rigorous, and should have met all the standards required for such research by scholarly, peer-reviewed professional journals. Gurin's report has not previously appeared in any such journal, nor could it be, for it fails to meet even the minimal standards for a peer-reviewed scholarly research report in a number of respects, including the failure to report how she codes her variables or how she selected her variables from those of the much larger HERI-CIRP 1985-89 undergraduate longitudinal study, and also her failure to justify the exclusion of two-year colleges from her sample. As such, her "Expert Witness Report" is not scholarly, peer-reviewed research. It is what is called "advocacy research."

The National Association of Scholars asserted in the brief that it filed in *Gratz v. Bollinger* that "[Neither the AALS nor the ACE brief] claims that campus racial diversity (much less campus racial diversity achieved through racial preferences) improves academic outcomes." The NAS also pointed out that the "University's amici ignore that portion of Astin's work that directly addresses the question at hand" (NAS Br. at 7)– i.e. the portion that finds no correlation between racial diversity and educational outcomes.

The Hon. Patrick J. Duggan noted in his opinion and ruling in *Gratz v. Bollinger* on 13 December 2000 that the NAS specifically took issue with the studies relied upon by the American Association of Law

[41] Not that this is an excuse for HERI's failure to release the database widely to the larger research community for more general research purposes.

Schools, contending that such studies really report that 'outcomes are generally not affected' by racial diversity on campus (NAS Br. at 6-7)."[42] *Significantly, Judge Duggan never disputes this claim in his opinion.* This is quite inexplicable on the merits, since the University's defense collapses completely if outcomes are *not* affected by racial diversity on the campus. And if it is true that the only existing national database that can adequately address this question actually *disconfirms* the University's claims, on what basis did Judge Duggan reach the conclusion that "... a racially and ethnically diverse student body produces significant educational benefits such that diversity, in the context of higher education, constitutes a compelling governmental interest under strict scrutiny"?[43] The lapse in his argument, it seems to us, can only be explained if one assumes that Judge Duggan took the question whether diversity *activities* are connected with positive educational outcomes to be the same thing as the question whether the racial diversity of the student body is. But clearly, they are not the same thing.

In Step 3 of her regressions, Gurin uses the racial diversity of an institution as a control variable. This means, as she notes herself, that when a campus experience variable, such as the discussion of racial issues, is a statistically significant predictor of some final outcome (such as socialization with minorities after graduation), then, according to her model, the size of the effect is independent of the number of minorities on campus. Thus the claimed beneficial effects of her campus experience variables would remain statistically significant even if the number of minorities on campus were to drop.

Gurin never incorporates all of her campus experience variables into a single regression. For all of her rhetoric, all she is doing is running regressions that leave out relevant explanatory variables. If she were to include all of her four diversity activities in one model, it is likely that the effect ascribed to the percentage of minorities on campus would disappear.

In the following, we develop some of these points further.

A Terminological Point

Since her Expert Report was written to defend the University of Michigan's racially preferential admissions policies, the key variable for her study was what Astin (1993c) calls the racial composition of the peer group, or what Gurin calls the "structural diversity" of the institution. Her task was to demonstrate that structural diversity has

[42] Duggan Opinion: 25.
[43] Ibid., p. 26.

direct effects, not indirect ones.[44] This she fails to do, just as Astin did. Like Astin and Chang, she resorts instead to what we have called the "in turn" hypothesis:

> Institutions of higher education that deliberately provide opportunities for positive intergroup interactions as they improve the representation of different racial/ethnic groups on campus are able to create the conditions for the positive effects of diversity on student development. As the educational institution becomes more multicultural in focus and its functioning, it is able to realize the benefits of various forms of diversity for all students. Research supports these different points and show that *structural diversity improves opportunities for interaction, which in turn, has positive effects on learning and democracy outcomes.* [italics in original][45]

Before entering into a more detailed criticism of Gurin's Expert Report, it is necessary to make a terminological point about her intermediate "interaction" variables.

In Gurin's analysis, special attention is paid to whether or not a student takes an ethnic studies course. Gurin calls this variable "classroom diversity." In our view, this is likely to be misleading, since it invites confusion with the very different concept of "racial diversity in the classroom." The latter concept would be the classroom counterpart of the campus variable "structural diversity," or the racial composition of the student body. It is important to note that there is no necessary connection between Gurin's "classroom diversity" variable and racial diversity, since it is entirely possible for a white student, for example, to take an ethnic studies course that has no blacks and Latinos in it at all, and vice versa. Accordingly, we believe it is preferable to label this variable "Took Ethnic Studies Course," and we will continue to do so throughout the ensuing discussion.

---

[44] Astin has defined "direct" and "indirect" effects as follows (1993c: 313): "[Direct environmental effects] are those that are unique to the environment in question and cannot be attributed to other environmental variables. What do we mean when we say that a particular environmental characteristic has a 'direct' or 'indirect' effect? A variable can be considered as having a direct effect when it enters the regression equation and maintains a significant Beta coefficient even after all other variables have entered the equation. … When this happens, the environmental variable is continuing to make a *unique* contribution to the outcome that cannot be accounted for or explained entirely by the effects of other environmental variables. An indirect effect is said to occur when (1) an environmental variable has a significant Beta coefficient after inputs have been controlled, but (2) the coefficient shrinks to nonsignificance when other environmental variables are added to the equation. In other words, when the effect of a particular environmental variable can be completely explained in terms of other 'mediating' variables, then its effect on the outcome is said to be entirely indirect."

[45] Gurin 1998: B, 2.

We give below the wording of the survey items from the 1989 HERI undergraduate survey on which Gurin's campus experience variables are based.[46]

> Since entering college have you (Yes/No):
> Enrolled in an ethnic studies course
> Attended a racial/cultural awareness workshop
>
> For the activities listed below, please indicate how often– Frequently, Occasionally, or Not at all– you engaged in each during the past year:
> Discussed racial/ethnic issues
> Socialized with someone of another racial/ethnic group[47]

The Relationships That Gurin Claims Are Statistically Significant Are Exceedingly Weak

The only part of Gurin's analysis that can shed any light at all on this question is Step 3 of her models. An examination of this regression step shows why Gurin never discusses the correlations between structural diversity and the student outcome variables–the only correlations that, in the final analysis, are relevant to the issue before the court.

Gurin's Expert Report presents four different multiple regression models for exploring the impact of various factors (which she calls "campus experience variables") on 16 student learning and 7 democracy outcomes. Each model includes whether or not a student attended an ethnic studies class (Gurin calls this "classroom diversity"). In addition, each model includes one of the following four variables:

- Attending a racial/cultural awareness workshop
- Discussion of racial issues
- Socializing with someone from a different racial/ethnic group
- The proportion of close friends in college who were of a different (or same) race/ethnicity

Her statistical analysis proceeds in three stages or steps. In the first step she controls for "student background characteristics" (such as SAT score, high school GPA, ethnic diversity of high school classmates, etc.), in order to determine the impact that these factors have on student learning and democracy outcomes. In the second step she adds the campus experience variables to investigate how they, in

[46] One of these variables apparently comes from a 9-year HERI longitudinal survey that Gurin and Dey incorporated into the Gurin Report, since the corresponding item is not found in the 1989 FUS (see Wingard et al. 1989: 201-208). So far as we know, the instrument used in this 9-year study has not been published, nor has HERI issued a comprehensive report of the findings. Few details about this part of the database are provided in the Gurin Report. As a result, we are unable to give the exact wording of Gurin's fifth "campus experience variable."

[47] From the 1989 HERI-CIRP Follow-Up Survey, Wingard et al. 1989: 204, 206.

combination with student background characteristics, impact student outcomes.

According to the model that she uses, there is a statistically significant but very small positive effect on a number of student outcome variables for white students who take ethnic studies courses and engage in other activities--workshops, discussion, etc.. She finds far less evidence of a positive impact for college experience variables among black and Hispanic students, even though she uses a less stringent test for statistical significance (.10 rather than the generally accepted .05 probability that a result occurs by chance).

In any case, both the Step 1 and Step 2 models are completely irrelevant to the question before the court in the University of Michigan litigation. All these models show is that taking an ethnic studies course and/or attending a workshop, discussing racial issues, etc. have a small positive impact on some student outcome variables. If the question before the court were whether the University of Michigan should be allowed to continue to have ethnic studies classes, sponsor racial/cultural awareness workshops; or ban inter-racial socialization or the discussion of racial issues, these results might be of some interest. But these are not the questions before the court, nor will they be before the court in any other litigation over racially preferential university admissions policies.

The issue is simply whether the University of Michigan is permitted by the federal constitution to manipulate the proportion of non-white students attending the university by explicitly taking a student's race into consideration during the admissions process. The only part of Gurin's analysis that can shed any light at all on this question is Step 3 of her models. In Step 3, Gurin adds variables to control for the impact of "institutional characteristics." These include the type of school (college or university), control (public or private), selectivity (mean SAT), faculty and institutional emphasis on diversity, and the one variable of most interest in answering the question posed in the previous paragraph: structural diversity, or the percentage of undergraduate students who are African American, Hispanic or Native American. This is exactly the variable the University of Michigan seeks to continue to manipulate by taking race into consideration during the admissions process. Step 3 of Gurin's models reveals what impact, if any, the proportion of "students of color" has on the correlations between the campus experience variables and final student outcomes. An examination of the nature of these relationships reveals why Gurin never discusses the correlations between structural diversity and the student outcome variables–the only correlations that, in the final analysis, are relevant to the issue before the court.

Nine-year outcomes

Gurin reports a statistically significant correlation between the percentage of minority students on a campus and participation and discussion of racial and ethnic issues nine years after graduation. The issue is how well PCTALL–the percent of all students who are minorities–predicts Discuss-Race. This impact is directly proportional to the size of the regression coefficient of PCTALL.

Instead of just controlling for all relevant student and campus characteristics, Gurin gives a dozen regressions for each race that rely on somewhat different sets of explanatory variables. In the 12 models where the sampled population consists only of white students, the coefficients of PCTALL are respectively: 0.0531, 0.0440, 0.0033, 0.0532, 0.0237, 0.00245, 0.0532, 0.0247, 0.0019, 0.0535, 0.0392, and 0.0029. The smallest coefficients, those around 0.003 (as opposed to those ranging from 0.02 to 0.05) occur in the four models in which institutional characteristics and those of students are subject to the most controls. These are the most inclusive models and the ones that best isolate the effect of PCTALL by controlling for other relevant factors. The first of the four models includes participation in an ethnic studies workshop. The second drops participation and uses campus discussion of racial issues instead. The third substitutes socialization with minorities while on campus as a control variable. The fourth uses racial diversity of close friends in college. Thus each of these four models involves 13 variables, 12 of which occur in all three.

The impact of coefficients of this magnitude depends on how the variables were coded. The relevant question on the CIRP survey was "Please indicate how often (**F**requently, **O**ccasionally, or **N**ot at all) you discussed racial/ethnic issues." Responses involved choosing **F**, **O** or **N**. In order to regress these responses on the 12 or 13 input variables, the responses must be coded numerically. In the absence of any explanation to the contrary, we would assume that "frequently" was coded as 2, "occasionally" as 1, and "not at all" as 0. The predicted value of the variable "discuss race" after controlling for 12 other variables is thus a score on an index that varies continuously from 0.000 to 2.000.

Gurin's model predicts that if the proportion of minorities goes up from 10 percent to 15 percent of all students, then the index of the variable - Discuss-Race would go up by 0.003 times 5 percent, which is an increase of only 0.015. More exactly, the predicted index increases for the four models are 0.0165, 0.012, 0.0095 and 0.0145%.

The highest of her three P-values (i.e, the one that is least significant) is 0.03, but this just means that one can assert with 97 percent confidence that the true value of the regression coefficient in this model

is greater than 0. Using her data one can compute a 95 percent confidence interval for the sizes of the coefficients she estimates. The largest of these coefficients is 0.0033 with a standard error of 0.000892. It follows that a 95 percent confidence interval for the true value of this coefficient in this model is 0.0033 ± 0.00175 [0.0033 plus or minus 0.00175] Even if one used the upper bound of this confidence interval as an estimate for the true regression coefficient, a five percent increase in the proportion of minorities on campus translates only to an increase of 0.025 in the index value for discussing race 9 years after graduation.

Let us take another example. One of Gurin's 9 year "democracy" outcomes is post-college-socialization with members of other racial/ethnic groups (which we will call PCSOCIAL). In one of her models she uses as a predictor socialization-with-other-groups while in college (SOCIALIZ). The reported regression coefficient for whites is 0.254. The precise meaning of this coefficient again depends on how responses were coded. The question asked was how often the respondent socialized with someone of another racial/ethnic group. The possible responses were frequently, occasionally or never. The obvious way to code responses would be 2 for frequently, 1 for occasionally, and 0 for never. We can then conclude that P–the index for socialization after college– is 0.254 higher for the group that socialized frequently than for the group that socialized occasionally while in college. This index is also 0.254 higher for the group that socialized occasionally in college than for the group that never socialized with members of other groups. (It is an artifact of the model that these two numbers are equal.)

But the coefficients of PCTALL in Gurin's four most complete models are 0.00933, 0.00896, 0.00647, and 0.00791. This means that a change from 5 percent minority to 15 percent minority produces a predicted change of at most 0.09 in this index. This is scarcely a third of the change associated with moving from the never socialized while in college and the occasionally socialized while in college groups, and scarcely one-sixth of the change associated with moving from the never to the frequently socialized groups.

Four-year learning and democracy outcomes

We continue our discussion of Step 3 of Gurin's modeling–which reveals what impact, if any, the proportion of "students of color" has on the correlations between campus experience variables and final student outcomes–by looking at the four-year learning and democracy outcomes for white, black, and Hispanic students. Table 1 summarizes the relationships Gurin's analysis revealed between the proportion of students of color and the 23 learning and democracy outcomes for white students. Tables 2 and 3 do the same for black and Hispanic

students. If the analysis revealed no statistically significant relationship it is coded as “NONE.” If the analysis revealed a statistically significant positive relationship (at the 0.05 level), it is coded as “Positive.” If a statistically significant negative relationship was revealed, it is coded as “Negative”.

Table 1

Impact of increased structural diversity on student outcomes, white students

Learning Outcome Variables	Workshop	Discussion	Socialization	Close Friends
College GPA	NONE	NONE	NONE	NONE
Highest Degree Desired	NONE	NONE	NONE	NONE
Academic Ability (vs Peers)	Positive	Positive	Positive	Positive
Drive to Achieve	NONE	NONE	NONE	NONE
Self Confidence (Intellectual)	NONE	NONE	NONE	NONE
Writing Ability	NONE	NONE	NONE	NONE
Listening Ability	Positive	Positive	Positive	Positive
Write Original Works	NONE	NONE	NONE	NONE
Create Original Art	Positive	Positive	Positive	Positive
General Knowledge	Positive	Positive	NONE	Positive
Problem Solving Skills	NONE	NONE	NONE	NONE
Critical Thinking Skills	NONE	NONE	NONE	NONE
Writing Skills	NONE	NONE	NONE	NONE
Foreign Language Skills	NONE	NONE	NONE	NONE
Preparation for Grad/Pro School	NONE	NONE	NONE	NONE
Democracy Outcome Variables				
Influence Political Structure	NONE	NONE	NONE	NONE
Influence Social Values	NONE	NONE	NONE	NONE
Help Others in Difficulty	NONE	NONE	NONE	NONE
Clean Up Environment	NONE	NONE	NONE	NONE
Participate in Community Action	NONE	NONE	NONE	NONE
Promote Racial Understanding	NONE	NONE	NONE	NONE
Cultural Awareness	NONE	NONE	NONE	NONE
Acceptance of Different Races	NONE	NONE	NONE	NONE

Table 2

Impact of increased structural diversity on student outcomes, black students

Learning Outcome Variables	Workshop	Discussion	Socialization	Close Friends
College GPA	NONE	NONE	NONE	NONE
Highest Degree Desired	NONE	NONE	NONE	NONE
Academic Ability (vs Peers)	Negative	Negative	Negative	Negative
Drive to Achieve	NONE	Negative	Negative	Negative
Self Confidence (Intellectual)	NONE	NONE	NONE	NONE
Writing Ability	Negative	Negative	Negative	Negative
Listening Ability	NONE	NONE	NONE	NONE
Write Original Works	NONE	NONE	NONE	NONE
Create Original Art	NONE	NONE	NONE	NONE
General Knowledge	NONE	NONE	NONE	NONE
Problem Solving Skills	NONE	NONE	NONE	NONE
Critical Thinking Skills	NONE	NONE	NONE	NONE
Writing Skills	NONE	Negative	Negative	NONE
Foreign Language Skills	Negative	Negative	Negative	Negative
Preparation for Grad/Pro School	NONE	NONE	NONE	NONE
Democracy Outcome Variables				
Influence Political Structure	NONE	NONE	NONE	NONE
Influence Social Values	NONE	NONE	NONE	NONE
Help Others in Difficulty	NONE	NONE	NONE	NONE
Clean Up Environment	NONE	NONE	NONE	NONE
Participate in Community Action	NONE	NONE	NONE	NONE
Promote Racial Understanding	NONE	NONE	NONE	NONE
Cultural Awareness	NONE	NONE	NONE	NONE
Acceptance of Different Races	NONE	NONE	NONE	NONE

Table 3

Impact of increased structural diversity on student outcomes, Hispanic students

Learning Outcome Variables	Workshop	Discussion	Socialization	Close Friends
College GPA	NONE	NONE	NONE	Negative
Highest Degree Earned	NONE	NONE	NONE	NONE
Highest Degree Desired	NONE	NONE	NONE	NONE
Academic Ability (vs Peers)	NONE	NONE	NONE	NONE
Drive to Achieve	NONE	NONE	NONE	NONE
Self Confidence (Intellectual)	NONE	NONE	NONE	NONE
Writing Ability	Positive	Positive	Positive	Positive
Listening Ability	NONE	NONE	NONE	NONE
Write Original Works	NONE	NONE	NONE	NONE
Create Original Art	NONE	NONE	NONE	NONE
General Knowledge	NONE	NONE	NONE	NONE
Problem Solving Skills	NONE	NONE	NONE	NONE
Critical Thinking Skills	Positive	Positive	Positive	Positive
Writing Skills	Positive	Positive	Positive	Positive
Foreign Language Skills	NONE	NONE	NONE	NONE
Preparation for Grad/Pro School	NONE	NONE	NONE	NONE
Democracy Outcome Variables				
Influence Political Structure	NONE	Positive	NONE	NONE
Influence Social Values	NONE	NONE	NONE	NONE
Help Others in Difficulty	NONE	NONE	NONE	NONE
Clean Up Environment	Positive	Positive	Positive	Positive
Participate in Community Action	NONE	NONE	NONE	NONE
Promote Racial Understanding	NONE	NONE	NONE	NONE
Cultural Awareness	NONE	NONE	NONE	NONE
Acceptance of Different Races	NONE	NONE	NONE	NONE

Tables 1-3 make clear that for the vast majority of student outcome measures, Gurin found no statistically significant relationship between structural diversity and campus experience variables, on the one hand, and final outcome measures on the other. Only 50 of the 276 models listed above show any statistically significant impact of increased structural diversity on student outcome variables. Of these 50 statistically significant relationships , over one third (i.e., 18) revealed negative effects of increased structural diversity on student outcome variables, and the vast majority of the negative impacts - 17 of 18 - were on black student outcomes.

Table 4 below summarizes the relationships between increased structural diversity and student outcomes from the previous four tables. It divides student outcomes into groups based on how increased structural diversity effects that particular campus experience variable. The results are: 9 Positive, 6 Negative, and the rest NONE. The table shows that increased structural diversity does not have consistent effects either in direction or across groups.

The first group of variables listed are those in which increased structural diversity is associated with better outcomes for at least one

Table 4: Summary of impact of increased structural diversity on student outcomes

POSITIVE	GAIN	LOSS
Listening Ability	White	None
Create Original Art	White	None
General Knowledge	White	None
Critical Thinking Skills	Hispanic	None
Clean Up Environment	Hispanic	None
Influence Political Structure(?)	Hispanic	None
MIXED	GAIN	LOSS
Academic Ability (vs. Peers)	White	Black
Writing Ability	Hispanic	Black
Writing Skills	Hispanic	Black
NO EFFECT	GAIN	LOSS
Highest Degree Desired	None	None
Write Original Works	None	None
Problem Solving Skills	None	None
Critical Thinking Skills	None	None
Preparation for Grad/Pro School	None	None
Influence Social Values	None	None
Help Others in Difficulty	None	None
Participate in Community Action	None	None
Promote Racial Understanding	None	None
Cultural Awareness	None	None
Acceptance of Different Races	None	None
NEGATIVE	GAIN	LOSS
Drive to Achieve	None	Black
Foreign Language Skills	None	Black
College GPA	None	Hispanic

group of students. The students who benefit are listed by race in the "Gain" column. The next group of variables is the mixed category, in which one group gains from increased structural diversity and another loses. The groups helped and hurt are listed in the "Gain" and "Loss" columns respectively. The next (and largest) group of variables are those in which there is no relationship between increased structural diversity and student performance for any group. Finally, the "Negative" group lists those variables where increased structural diversity hurts student outcomes. Those hurt are listed in the "Loss" column.

Table 4 shows that for 14 of 23 student outcome variables, increased structural diversity has either a negative impact on one or more groups or no effect at all. For three of the remaining nine variables, the impact is mixed. For only six of the 23 outcome variables is there a positive effect on either white or Hispanic student performance with no decline in the performance of students of another group

Simply stating that there is a statistically significant positive effect for a small number of variables for some groups is insufficient. In order to assess the impact of increased structural diversity on student outcomes, it is also necessary to know how large (or small) the effects are of increased structural diversity on student outcome measures.

Although Gurin never gives the magnitude of the relationships her analysis discovered, it is easy to compute these magnitudes because her regression analysis predicts the impact that changes in structural diversity and other variables of interest (such as taking an ethnic studies class) would have on student outcomes.

Table 5
Regression coefficients, t-statistics, and significance levels
for white student four-year listening ability outcomes

Listening Ability Independent Variables	4 Point Scale Coefficient	t-Stat	Sig T
Ethnic Studies	0.0761	3.591	0.0003
Workshop	0.1417	6.354	0
Percent Minority	0.0058	4.442	0
Ethnic Studies	0.0693	3.267	0.0011
Discussion	0.1173	7.28	0
Percent Minority	0.0054	4.123	0
Ethnic Studies	0.0847	4.069	0
Socialize	0.1172	7.72	0
Percent Minority	0.0043	3.229	0.0012
Ethnic Studies	0.1004	4.849	0
Friends	-0.0375	-2.817	0.0049
Percent Minority	0.00559	4.182	0

Let us pick the best case for Gurin (by her own standards): one where there is a statistically significant positive relationship in *all four* of her models between taking an ethnic studies course (what she calls "classroom diversity"), three of four "campus experience" variables, and increased structural diversity and an outcome variable. One example is the four-year listening ability outcome for white students.[48,49]

Table 5 lists the regression coefficients, t-statistics and significance values for white students for the four-year listening ability outcome variable from Gurin's regression analysis. Any coefficient with a t-statistic with an absolute value greater than or equal to 1.96 is considered statistically significant. The probability that such a coefficient is actually zero is less than 0.05, or 5 percent. The larger the t-statistic the lower the probability that the actual value of the coefficient is zero. The Sig T column lists the probability that each coefficient is zero based on its t-statistic. In Table 5 all of the coefficients have t-statistics with absolute values greater than 1.96, and Sig T values of 0.05 or less. So, all of these coefficients are considered to be statistically significant. But what do they mean in practical terms?

[48] See line 7 of Table 1 above and Gurin 1998: Table D.2 p.2 of 12.
[49] There were no instances where all campus experience variables and the structural diversity variable were significant and positive for any four-year outcome variable.

Let's look at the workshop model. Imagine that a university offers a voluntary ethnic studies class and a voluntary ethnic workshop. Students attending this university can engage in various levels of discussion of racial issues, and socialize and make friends with people of different races at different rates. Assume that 10 percent of the university's students are minorities. Gurin's "workshop" regression analysis tells us that, on average, if a student at this university decides to take an ethnic studies class, the student's listening ability score (on a scale of 0 to 4) rises by 0.0761, or about two percent. If the same student also takes the workshop, the model tells us that his or her score rises by 0.1417, or about 3.5 percent. These effects are additive, so if a student engages in both activities, he or she can expect about a 5.5 percent increase in listening ability.[50] In addition, Gurin's model predicts that for each additional one percent of the student body that consists of minorities, a white student's listening ability will improve by 0.0058 on the four point scale, or about 0.145 percent; i.e., 14 hundredths of one percent. So for our college where 10 percent of the student body is minority, the average white student's listening ability would improve by 10 times 0.145 = 1.45 percent compared to a hypothetical college with an entirely white student body. A white student at our hypothetical university who takes both the ethnic studies class and the workshop can expect to have listening skills about 7 percent better than a student who took neither the class nor the workshop at a hypothetical school with no minorities. These relationships hold regardless of changes in any of the variables Gurin controlled for in her model, such as SAT score, high school GPA, student's gender, selectivity of the school, etc.[51] These relationships hold true regardless of the levels of the variables.

For 14 of Gurin's 23 student outcome variables, increased racial diversity has either a negative impact on one or more groups or no effect at all. For three of the remaining nine variables the impact is mixed. For only six of the 23 outcome variables is there a positive effect on either white or Hispanic student performance with no decline in the performance of students of another group.

Thus, if a student took the ethnic studies class, but not the workshop, his or her listening ability would still improve by about two percent. Similarly, if the percentage of the student body that is minority falls from 10 to 5 percent, taking a workshop would still improve the student's listening ability by 3.5 percent. In other words, Gurin's analysis shows that, in those cases where "Took an Ethnic Studies

[50] Contrary to the impression Gurin gives in her report, it is the small size of the effects she is looking for, and not inadequate sample size, that limits her ability to detect effects for black and Hispanic students in the CIRP data base. Although Gurin tells us several times that the CIRP data base she used contained data on 9,316 students, she does not tell us the proportion of black and Hispanic students in the samples she used. The conservative assumption that at least 5 percent of the students in the overall data base were black and 5 percent Hispanic would give Gurin a sample size of 465 for each group. This is larger than the samples used in much social science research and in any case is large enough to detect effects of any practical significance for purposes of public policy making.

[51] See Gurin Appendix C pages 13-14 for a complete list of the variables she controlled for in her models.

Course" and the other "campus experience" variables improve student outcomes, *their effects are independent of the fraction of students that are minority.*[52] In fact, by structuring her analysis as she did, she explicitly assumes that structural diversity and campus experience variables have *independent* effects. The results she calculated for her campus experience variables were deliberately calculated to be valid at any level of structural diversity.

We have previously observed that most of Gurin's variables are soft student self-assessments. Students' self-evaluated listening ability is so vaguely defined as to make the small changes predicted by a model especially meaningless. Student attitudes that cause them to take these courses or to participate in such workshops are probably more responsible for any apparent effects then the courses or workshops themselves. Some evidence for this claim can be found in the relative magnitudes of Gurin's reported coefficients, which, if taken seriously, would show that the impact of a single workshop (involving at best a couple of hours) is almost twice that of taking a whole course. More specifically, the coefficient for taking a workshop is 0.14, that for taking an ethnic studies course for an entire semester is 0.076.

Gurin's 1985 CIRP variables do not sufficiently control for political and other attitudes that may predispose students to take certain courses, or to engage in certain activities. The 1994 CIRP questions (on which Chang relies) provide much better controls for characteristics that students bring to college, or which predispose them to chose a certain college. The Michigan survey ("The Undergraduate Experience at Michigan") does include numerous items like religious affiliation, whether the respondent is religious, campus activities (including religious, fraternity, student government, and athletic participation), hours worked, personal and political attitudes, and many others. But, of course, data at any one institution cannot be used to estimate the effects that may arise from varying percentages of minorities on different campuses.

Gurin's analysis shows that even when compared to the modest impact of campus experience variables on student outcomes, the impact of structural diversity as measured by the percentage of minority students is very small. To take the example just given, suppose that the

Suppose that a university decides to stop taking race into consideration in its admission practices, and that the fraction of non-white - Asian students falls by 50 percent from 10 percent to 5 percent. What impact would this have on the (self-reported) white student listening ability outcome? Gurin's results predict that white student listening ability would decline at this institution by 5 x 0.145 = 0.725 percent, or about seven-tenths of one percent.

[52] These conclusions, which are the only valid reading of her regression analysis, directly contradict the assertions Gurin makes about the role of structural diversity between pages 31 and 35 of her report. One can only assume the correlations presented in Table 1 on page 32 and the data presented in Figure 2 do not control for the student background and institutional characteristics variables listed on pages 13 and 14 of Appendix C of Gurin's Expert Witness Report. Therefore, these earlier assertions are subject to a variety of biases and errors that are controlled for in the regression models, and should be regarded with great suspicion.

university decided to stop taking race explicitly into consideration in its admission practices, and that the fraction of minority students falls by 50 percent from 10 percent to 5 percent. What impact would this have on the (self-reported) white student listening ability outcome? On average, Gurin's results predict that white student listening ability would decline at this institution by 5 x 0.145 = 0.725 percent, or about seven-tenths of one percent.

The results for the "Discussed Race," "Interracial Socialization," and "Close college friends were diverse/same" models can be calculated in a similar way. In general, the coefficients are very similar in magnitude to the workshop model and the results will be very similar for the other three models. The notable exception is the "Close college friends were diverse/same" model, which predicts that for white students an increase in the proportion of friends from other racial/ethnic groups leads to a small *decrease* in listening ability. Results for other outcome variables can be calculated in a similar manner.

Thus, about 88 percent of Gurin's models show that there is either no relationship, or a negative relationship, between structural diversity and both learning and democracy outcomes. In those few cases where structural diversity does have a positive impact on learning or democracy outcome variables, its impact is extremely small, as illustrated by the example above.

In addition, her results show that any positive effects of campus experience variables such as ethnic studies courses and workshops are independent of structural diversity. In other words, the effectiveness of ethnic studies courses, workshops, etc. remains the same no matter what the percentage of students that are minority at a particular school, or how that percentage changes over time.

We know from Gurin's regression tables that the effectiveness of ethnic studies courses, workshops, etc. remain the same no matter what the percentage of non-white - Asian students is at a particular school, or how it changes over time.

From our examination of the 850 pages of regression tables we were able to obtain, our impression is that this is true for all the other student outcome variables in the Gurin/Dey CIRP database (i.e., besides the 15 learning and 8 democracy outcomes that are included in her tables). Take, for example, the College Satisfaction variable. This variable is of interest because it throws additional light on the survey findings we reported in Part III of this report. There we mentioned that the Zogby/FAST survey found that students are overwhelmingly opposed to racially preferential university admissions policies. On the other hand, defenders of preferences can cite this survey (and others) in support of the proposition that students value racial diversity, and that they believe that campus racial diversity is important to higher education. At this point, an outside referee might conclude that the debate is moot. This, however, would be a mistake, because the question is an empirical one that can be addressed in a perfectly

straightforward way through multivariate regression analysis, provided that the database includes measures of college satisfaction. As it turns out, the CIRP database includes this measure.[53]

We already know, from Astin 1993c: 362, that there is no educationally significant correlation between structural diversity and the College Satisfaction outcome variable. (If there had been, Astin would have reported it.) But the Gurin/Dey regression tables we obtained permit us to significantly strengthen this conclusion. In particular, they have enabled us to test the hypothesis that increased structural diversity will, in conjunction with what Gurin calls the "campus experience variables" (e.g., taking an ethnic studies course, discussing racial issues) have a positive impact on final student outcomes. The tests tell us to reject this hypothesis. We summarize the results in the following tables:

Overall Satisfaction: four-year study				
	Workshop	Discussion	Socialization	Close Friends
White	NONE	NONE	Negative	NONE
Black	NONE	NONE	NONE	NONE
Hispanic	NONE	NONE	NONE	NONE

Overall Satisfaction: nine-year study				
	Workshop	Discussion	Socialization	Close Friends
White	NONE	NONE	Negative	NONE
Black	NONE	NONE	NONE	NONE
Hispanic	NONE	NONE	NONE	NONE

In sum, Gurin wants the University of Michigan to continue to trample on the Fourteenth Amendment rights of every applicant by explicitly considering his or her race in its admission decision in order to increase structural diversity, even though her own analysis shows: (1) few instances of positive effects of structural diversity on either learning or democracy outcomes; (2) very small (and educationally insignificant) effects where the correlations are positive and statistically significant; and (3) no connection between structural diversity and "campus experience" variables such as ethnic studies courses, workshops, etc.

[53] See item 10 of the FUS, Wingard et al. 1989: 205.

Gurin's Expert Testimony Fails to Meet the Standards for Research Reports in the Peer-Reviewed Scholarly Literature

Gurin's "Expert Testimony" fails to report data that would be included in a scholarly research article

If we had access to the CIRP data set, we could report the estimated values themselves, not just the changes arising from different levels of the percentage of minority students. We could report, for example, something like the following: for white female students with a B+ high school grade point average, an SAT total of 1200, who had almost no minority friends in high school but who socialized with minorities in college, etc. (eight more variables), the index for discussing race after graduation is .820 if they went to a college with only 10 percent minorities, and is .835 if the proportion of minorities was 15%. We repeat that the 0.820 is a made-up number; we would need access to the data to know what number is actually predicted by this model. But the difference, according to Gurin's own model, that is made by a change from 10 percent to 15 percent minority is what we have calculated: 0.015.

The difference according to Gurin's models between a campus that is 5 percent minority–a combined proportion of blacks, Asians and Hispanics almost certain to be exceeded without racial preferences–and one with 25 percent minority is a change of 4 times that computed above; i.e., around 0.06 in the discussed race index–hardly a dramatic change.

This data would ordinarily be reported in a scholarly research report. But its omission is a particularly serious matter given the claims that Gurin has made, for while racial preferences can be used to increase racial diversity on a campus, some degree of racial diversity will normally be attainable in their absence.

Gurin fails to report the criteria she used in selecting the variables and samples of institutions she uses

The value of coefficients depends both on what other variables are included in a model and on the population being studied. For example, given a population of grade school boys one could predict weight from any of the three variables height, age or shoe size (ignoring width; i.e., coding size 8B as just 8). It would turn out that if one regressed weight on age alone (used age to predict weight), the coefficient for age would be highly significant. If one regressed on shoe size alone, the coefficient of shoe size would also be significant. Each of the three explanatory variables looked at by itself would have a statistically significant regression coefficient. However, if one combined the three

variables into one model, one would probably find that only the coefficient of height was significant, either statistically or practically. If one knows a boy's height, additional information on age and shoe size does not materially improve the accuracy with which one can predict his weight. Thus shoe size and age are significant by themselves, but not when one includes the more appropriate variable of height.

The effect of the population being studied can be seen if one looks at the same variables in a population of men in their 20's. One might then find that age is of no value in predicting weight, and that shoe size is a much weaker predictor than in a population of young boys. Of course, height would still be the best predictor.

Gurin recognizes that different populations can yield different regression coefficients when she analyzes data separately for each race. But this recognition adds force to her failure to explain or justify the selection of her data set, which consists of a relatively small part of the data previously analyzed by Astin. Gurin reports that she dropped historically black colleges from the analysis. If it turned out that– all else being equal–minority students do better at such schools, her case would be weakened. Gurin also dropped two-year institutions from the analysis, without explaining why.

If they were given access to the databases, researchers might be able to assess the impact of racially preferential admissions policies

If we had access to the CIRP data, we could pursue various threads suggesting that there are negative effects associated with a system of racial preferences. CIRP does not contain any pure measure of racial preference on campuses, but a reasonable measure might be the difference between the SAT scores of whites and blacks on a campus. If we had access to the CIRP data set, we might be able to utilize external data on the racial SAT gap at various schools, and then assess the impact of this measure on Gurin's outcome. The best proxy that we have in the existing data set for the degree of racial preference might be the selectivity of a school (measured by the average SAT scores of admitted student), because more selective schools tend to grant a greater degree of preference to minorities on the SAT. (Of course, many schools admit virtually all applicants with a high school diploma, but while these schools may put more effort into recruiting blacks than whites, they do not discriminate by race in admissions.)

Astin (1993c: 51) reports that there is a very substantial correlation (0.39) between selectivity and racial conflict. It is conceivable that the explanation has something to do with the tensions caused by racial preferences at the more selective schools. We do not claim that this

one correlation provides proof of this thesis. We are simply pointing out that only one side has had access to the data.

Release the database!

HERI was, in effect, a cooperating organization in the production of Gurin's expert testimony. This is not the first time that HERI has cooperated with like-minded organizations in the defense of racially discriminatory university admissions policies. HERI did so as well for the Mellon Foundation. This foundation used the CIRP database (and not just regression tables generated from the database) in the College & Beyond study that is the basis of William Bowen and Derek Bok's book *The Shape of the River.*[54] This renders any refusal on the part of the University of Michigan to release its database (or any future refusal by HERI to release the entire database) even more untenable.[55]

It is a disturbing departure from standard research practice for a database of this size and importance to have gone so long without having been made available to other researchers in the field, who would then be in a position to peer review studies and analyses based on it. The failure to release the database to the general research community is rendered even more inexcusable by the fact that HERI has already turned the database over to universities and foundations that it regards as ideologically sympathetic. This is even more outrageous than a blanket refusal to release the database to anyone. As it is, it appears that the database is being made available *only* to organizations that are deemed to be ideologically sympathetic. In short, there appears to be selective bias at work here.

For the reasons we have given, it is very hard for us to imagine how the defenders of racially preferential admissions policies can believe that the CIRP database, or any of the studies that have been based on it to date, can be at all helpful to their cause. We would not be surprised to see universities and the higher education establishment at One Dupont Circle simply drop references to the CIRP findings in the future. If they do not, however, courts, the research community, and the general public must insist that the full database be made available to the entire

[54] Bowen and Bok 1998: 94n8, 147, 294, 313, 338.

[55] "As with the data from the College Entrance Examination Board, records were matched by HERI using an algorithm that protected the confidentiality of the records." (Bowen and Bok 1998: 314.) This means that one can dismiss in advance any concerns that HERI might express about maintaining the confidentiality of its database, since HERI has already found an algorithm that it feels provides adequate confidentiality. There are in fact standard methods and algorithms for the generation of "statistical databases" (Denning and Schlörer 1983, Adam and Wortmann 1989, Tendick and Matloff 1994).

research community, with no questions asked about the ideological orientation of the researchers who might wish to have access to it.

Gurin's Assertion That Statistical Significance Is a "Basic Indicator of the Strength" of Her Relationships Is a Fundamental—and Rather Elementary—Statistical Error

Gurin misleadingly or carelessly asserts (Appendix C, p. 2) that a "basic indicator of the strength of these relationships with the outcome measures is found in the assessment of its statistical significance." On the contrary, statistical significance is not a measure of the strength of a relationship; it is instead a measure of the probability that an effect of the apparent size could arise purely from chance variation in the selection of the sample. Statistical significance depends both on the intrinsic strength of an alleged effect and also on the sample size.

For example, suppose one took a poll and found that when asked a certain question, 62 percent of men and 60 percent of women said "yes." Most people would interpret this as meaning that there was no evidence of any meaningful sex difference. While the difference between 62 percent and 60 percent is insignificant in most cases, it can be *statistically* significant if the samples are large enough.

If the poll had used samples of 10,000 men and 10,000 women, the observed difference of 2 percent would be highly significant statistically, with a P-value of 0.0039 (i.e., 39 hundredths of one percent). This does not mean that the difference of 2 percent is important, but only that it cannot be explained by the luck of the draw, i.e., by who happened to be included in the sample. In this case one has reliable evidence of the existence of a small but real sex difference in responses.

But if the sample sizes were 100 men and 100 women, the P-value would be 0.77. This means that even if the true proportions of all men and all women ("all" means the entire population, not just those who happen to be included in the samples) who would say "yes" are identical, there is a 77 percent probability of choosing samples in which the observed difference will be 2 percent or more. The large P-value means that there is considerable doubt as to whether there is any sex difference at all.

The distinction between practical and statistical significance, between the size of an effect and the probability that an apparent effect is due to chance, is a standard topic in elementary statistics courses. One well known text explains the difference as follows:

> If a difference is statistically significant, then it is hard to explain away as a chance variation. But in this technical phrase, 'significant' does not mean 'important.' Statistical and practical significance are two different ideas. . . .[56]
>
> The P-value of a test depends on the sample size. With a large sample, even a small difference can be statistically significant, that is, hard to explain by the luck of the draw. This doesn't necessarily make it important. Conversely, an important difference may not be statistically significant if the sample is too small."[57]

Gurin's Models Fail to Control Simultaneously for All Student Characteristics

The sheer size of Gurin's analysis–850 pages of statistical printouts– may convey the impression that she has left no stone unturned. But her procedure is really designed to obscure the fact that the percentage of racial minorities on a campus has few if any discernible beneficial effects on her outcome measures, and probably no statistically significant effects at all. Gurin's obvious procedure would be to run a single regression for each of the outcome variables against all 16 of her explanatory variables. That is, instead of having 12 regressions (12 models) for each of her outcome variables for each of the three races, she would have only one for each race. Instead of 850 pages of printouts she would have only 70 pages.[58]

Instead she first regresses an outcome variable against five pre-college variables ("student characteristics"): high school grades, SAT, sex (female coded as 1, male as 0, we think), and two pre-college ethnic variables (proportion of high school classmates and of neighbors who were of the same race or ethnicity). She then appears to have looked at the coefficient for each of 8 additional variables if added singly (without any of the others) to the list of predictors. This procedure would be justified (as a first step in an exploratory data analysis) if her intention had been to discover which of these potential explanatory variables would be used in the final analyses. But since she doesn't eliminate any variables, she merely produces printouts for

---

[56] The authors then describe a hypothetical example–similar to the one given above– in which the difference between two groups is statistically significant because of the large sample size, but too small to have practical consequences.

[57] Freedman, D. , R. Pisani and R. Purves 1998: 554.

[58] If she wanted to facilitate evaluation of her claims, she would also have included other statistics, such as R^2– the percent of variance explained by the model, which is part of the standard output of regression programs. She would also have included summary statistics for her input measures that would, among other consequences, enable others to be sure how her variables were coded.

models in which some explanatory variables are omitted. If an omitted variable is correlated with one or more variables included in the regression, the omission will have the effect of falsely ascribing effects to included variables that should really be ascribed to omitted ones.

In a second series of models she adds taking an ethnic studies course and one other variable to the five and computes the coefficients for this model, after which she adds each of the remaining variables singly (without any of the others). Again this only makes sense if she intended to start a forward regression whose purpose was to select which variables would be used in the final analyses.

Finally, she regresses her outcomes against all but three of her variables, varying the omitted ones. This procedure has no statistical justification, in spite of her rhetorical flourishes in which she gives names to the various models (workshop, discussion, socializing, and racial diversity of close friends). All she is doing is leaving out relevant variables.

Why does she resort to these convoluted procedures? The most likely explanation is that when she experimentally included all 16 explanatory variables in a single regression, the coefficient for percent minority (PCTALL) was never (or practically never) statistically significant. She therefore decided to use models that utilize only one of the four variables, (workshop, discussion, socializing, close college friends are diverse) at a time. The apparent effect that she ascribes to PCTALL is most likely an artifact that arises from the omission in each of her models of at least three of these four variables.

Had Gurin run her regressions with all 16 explanatory variables simultaneously present, she would probably have duplicated Astin's results, which showed that PCTALL doesn't have any statistically significant beneficial effects on any of the outcome variables for members of any racial group. Had she used Astin's full data set and full set of variables, she would necessarily have duplicated his results.

Gurin claims that her results provide a

> conservative estimate of diversity's effects, in that the analyses consistently allow other variables in the analysis (i.e., characteristics of colleges and entering characteristics of students) a greater opportunity to account for, and possibly explain away the influence of diversity on college students. . . . Despite the fact that this approach tends to diminish the likelihood of demonstrating effects related to diversity, it is important to take these relationships into account in order unambiguously to demonstrate change related to diversity. In sum, this approach ensures that where I report significant

> effects related to diversity, they are truly diversity effects, as opposed to being consequences of the characteristics, choices and preferences that students bring with them to college.

Gurin claims to be generously stacking the deck against her own hypotheses. But her analysis would have had no credibility at all if she had made no effort to take into account student characteristics that predate their enrollment in college and characteristics of the institution other than the proportion of minority students. Furthermore, as stated above, her finding of significant (though small) diversity effects rests on her failure to use models in which all explanatory variables are simultaneously present.

A standard topic in elementary statistics courses is the difference between correlation (or association) and causation. To say that C causes D means that if one could change only C (without changing other relevant factors), then D would also change. A famous example, ironically cited during the prohibition era, was the high correlation between the price of rum in Havana and ministers' salaries in Boston. The source of the correlation was of course that both were being influenced by the same broad economic trends (prosperity and inflation). If, due to the normal workings of the economy, ministers' salaries rose, one could validly predict that the price of rum had gone up or would go up. But if a multimillionaire had intervened to award raises just to Boston clergy, the price of rum would not have been affected. Changing just this one variable while all else remained constant would not have affected the price of rum. While there is an association between these variables, the relationship is not causal.

Multiple regression models are often, though not always, able to disentangle these effects. Thus if one regressed rum prices on the average salaries of ministers, the regression coefficient would be not only statistically significant but large enough to enable one to make useful predictions. On the other hand, if one included economic variables on overall price levels, the coefficient for ministers' salaries would probably shrink to both practical and statistical insignificance.

Much of Gurin's analysis is impeached by a failure to account for obvious variables that would readily provide alternative explanations for her findings. In spite of Gurin's assurance that she is generously taking other explanations into account, she presents some tables which completely (as opposed to inadequately) fail to do so. Thus figure 2 (p. 33) of the Expert Report is labeled "structural diversity effects on interracial contact patterns after college among white students raised in predominantly white neighborhoods." According to the graph, for white students who attend the least diverse colleges (percent minority 0 to 9%), only 10 percent have diverse current friends, only 21 percent

have diverse current neighbors and only 24 percent have diverse current co-workers. However for students who attend the most diverse colleges (more than 25 percent minority), these percentages are respectively 26, 35, and 40. The size of these effects is vastly greater than those that Gurin claims to have found in her regressions, so one can be sure that the apparent effects are traceable to other student characteristics. Is it really plausible that recent graduates starting on careers and seeking places to live would make choices that are seriously influenced by the number of minority employees in a firm and the number of minority residents living in a neighborhood? Do those who seek jobs and apartments have so many choices that they can select jobs and apartments on the basis of these criteria as opposed to salary, opportunity for advancement, location, transportation, and rent?

Since Gurin's data has not been made widely available to the research community, it is impossible to verify or definitively determine what accounts for the numbers she reports, but plausible speculation still reveals the shakiness of her argument. Diversity of friends, neighbors, and coworkers are defined in a footnote as meaning that more than half were non-white–a rather unusual circumstance in most of the U.S. A student who attends a low diversity college would be more likely than most students to live in a state like Nebraska or Oregon or Wyoming, where there are relatively few minorities. Unless the student moves to a big city after graduation, he would find it hard to be in a situation where literally most of his friends, neighbors, and co-workers were non-white. On the other hand, those students who attend colleges that are more than 25 percent non-white probably also live in areas with sizable non-white populations. These students are likely to have minority friends, neighbors, or co-workers, just as a matter of local demographics, and not because they actively sought to do so and not because they went to a college that was 25 percent minority.

THE SUPPLEMENTAL EXPERT REPORT OF PATRICIA Y. GURIN (11 January 2001)

The National Association of Scholars raised some of the foregoing points in the amicus brief that it filed in the *Gratz v. Bollinger* ligitation.[59] In particular, the NAS pointed out there that the CIRP database found no educationally significant correlations between campus racial diversity and educational outcomes. We also pointed out that in the light of these findings, any claims about what Gurin calls "campus experience variables" are beside the point.

[59] National Association of Scholars (14 July 2000).

Gurin has replied to the NAS brief in a court document.[60] There Gurin has asserted that *we* have argued that the alleged benefits do not require the presence of minority students. She also protests that we accuse her of having treated enrollments in ethnic studies courses as an "adequate proxy" for racial diversity.

As to the first point: What the NAS actually argued was that the University clearly has the burden of showing whether, or to what extent, the four "campus experience variables" that Gurin studies require institutional campus diversity in order to function in the way she believes they do. Thus far the University has not produced this evidence, and we can infer from Gurin's reply to the NAS brief that it is in no position to do so. Indeed–and quite incredibly–the University has explicitly *refused* in court to specify the degree of racial diversity that it believes is required to produce the alleged educational benefits. This means, in effect, that we do not even know whether the *University of Michigan* believes that its racially discriminatory admissions policies are necessary to produce the alleged educational benefits!

As to the second point: Gurin has apparently misunderstood the underlying thrust of our argument, which has the form of a reductio ad absurdum. Our point has simply been that Gurin's findings about the "campus experience variables" could prove that the University's discriminatory policies meet the compelling state interest test in court only if these variables *were* an adequate proxy for campus racial diversity. Since they obviously cannot be, the whole argument fails. Now, however, we finally have a public admission from Gurin that her four campus experience variables cannot be treated as adequate proxies for the only input variable that could possibly be of interest to the court, i.e., the racial diversity of the institution. Her recent admission now raises in the sharpest possible form the following question: Why does Gurin treat the campus experience variables as the primary object of her study, rather than the racial diversity of the institution?

Nothing that Gurin says in her reply to the NAS' amicus brief even begins to address this question. After all, one does not have to go looking for "adequate proxies" in this case at all, whether it be interracial on-campus socializing, ethnic studies courses, or anything else, since the CIRP database which Gurin herself uses has not one but three variables that designate the very thing of interest–i.e., the racial diversity of the institution, or what Astin calls the racial composition of the peer group. And what we know from Astin and from subsequent studies at HERI is that one cannot assert that the racial diversity of a campus is correlated in any educationally significant way

[60] Gurin, P. (11 January 2001).

with positive educational outcomes when one properly controls for possibly confounding variables such as Gurin's four campus experience variables." And that, so far as the question before the courts is concerned, is the end of the matter.

Gurin also claims in the "Supplemental Expert Report" that her work "demonstrated a remarkable consistency in results."

However, since the results she finds are for variables that cannot, even by her own admission, be treated as adequate proxies for the one and only input variable that could possibly be of interest to the courts, this remark is entirely beside the point.

Here, too, the enthymeme of her argument has to be that her findings about the "remarkable consistency" of the impact of "campus experience variables" shows something indirectly about the impact of structural diversity on outcomes (although she never tells us what that is). But why all this indirection? It is pointless and misleading for the University to ask us to eyeball findings about variables that are not at issue in the litigation and to deem them "remarkably consistent," and then to ask us to make wholly unsupported and unsubstantiated inferences about the meaning of these correlations for the real variable of interest, when there is a *direct* test of the hypothesis in question which the University entirely ignores: indeed, it is a test that Astin had run nine years earlier in his comprehensive analysis of the same database!

In order to test the one hypothesis or claim that is of relevance to the court, one does not have to eyeball anything or make any decisions about when a pattern of findings is "remarkable" or not, since the CIRP database itself allows us to test the hypothesis that is at issue in the litigation *directly*: one simply tests for correlations between racial diversity and educational outcomes with models in which other explanatory variables are properly controlled. When one does that, however, it turns out that the findings tell us to *reject* the University's claims.

THE HON. PATRICK J. DUGGAN'S OPINION IN *GRATZ V. BOLLINGER*[61] (13 DECEMBER 2000)

On December 13, 2000, Judge Patrick J. Duggan of the U.S. District Court of Eastern Michigan issued a ruling and opinion upholding the University of Michigan's current race-based admissions policies. Perhaps the central contention on which this opinion and ruling was based was the following: "The University Defendants have presented

[61] Duggan Opinion.

this Court with solid evidence regarding the educational benefits that flow from a racially and ethnically diverse student body."[62]

Among the documents cited by Judge Duggan in support of this claim were briefs filed in the case by the United States Department of Justice, the Association of American Law Schools, and the American Council on Education:

> A number of amici have filed briefs concurring with the University that diversity results in a richer educational experience for all students. In support of its position, the United States cites a study by Alexander Astin, Director of the Higher Education Research Institute at the University of California, in which Astin associates diversity with increased satisfaction in most areas of the college experience and an increased commitment to promoting racial understanding and participation in cultural activities, leadership, and citizenship. (U.S. Br. at 20-21; *see also* ALS Br. at 6; ACE Br. at 15).[63]

The implication that the AALS, the ACE, and the United States have claimed that *Astin's* study supports the view that "diversity results in a richer educational experience for all students" is incorrect. A perusal of the passages cited by Duggan clearly shows that these amici were, if anything, very careful to *avoid* making this claim. Furthermore, we know from p. 362 of Astin's major study that they could *not* have made this claim (at least so far as Astin is concerned), because Astin found that the database tells us to *reject* this hypothesis.

The amicus briefs cited by Judge Duggan cite two publications by Astin: a 1993 article in *Change* magazine and *What Matters in College?* Since the magazine article is only a popularization of the findings that are fully reported only in the book, it suffices to quote the relevant passage from the latter:

> The study also included several indicators of the individual student's *direct experience with diversity activities* [emphasis ours]: taking women's or ethnic or Third World courses, participating in racial or cultural awareness workshops, discussing racial or ethnic issues, and socializing with someone from another racial or ethnic group. …One other student outcome that is positively associated with *individual diversity activities* [emphasis ours] …. is political liberalism. … In short, the weight of the empirical evidence shows that the actual effects on student development of *emphasizing diversity* [emphasis ours] and of

[62] Ibid., p. 21.
[63] Ibid., p. 22.

> *student participation in diversity activities* [emphasis ours] are overwhelmingly positive. …. [T]he findings of this study suggest that there are many developmental benefits that accrue to students when institutions encourage and support an *emphasis* [emphasis ours] on multiculturalism and diversity."[64]

Note that whereas Duggan attributes to Astin the view that racial *diversity* results in a richer educational experience for all students, Astin makes the quite different claim that "direct experience with diversity activities," " emphasizing diversity," and "student participation in diversity activities" are correlated with (some) positive educational outcomes. *Pace* Judge Duggan, the amici make only the latter claims, not the former one. [65]

That is why the National Association of Scholars asserted in the brief that it filed in the case that "Neither of these sources [i.e., neither the AALS nor the ACE brief] claims that campus racial diversity (much less campus racial diversity achieved through racial preferences) improves academic outcomes." The NAS also pointed out, in a passage that Judge Duggan does not cite, that the "University's amici ignore that portion of Astin's work that directly addresses the question at hand" (NAS Br. at 7)– i.e. the portion that finds no correlation between racial diversity and educational outcomes.

Perhaps Duggan confused the question whether diversity *experiences* are correlated with positive educational outcomes with the very different question whether the racial composition of the student body is so correlated. It is only research addressing the *latter* question that can answer the empirical question that is engaged by the present litigation. It only stands to reason that what Astin calls "diversity activities" are at best very weakly correlated with the racial diversity of the student body. But one need not rely on a priori considerations here. We *know* that racial diversity and these campus experience variables are imperfectly correlated, even within the very database that is at issue here, because, while Astin found that diversity activities are correlated in the database with at least some positive outcomes, he also found that the crucial variable–the racial composition of the student body–is not.

---

[64] Astin 1993c: 431.

[65] E.g., the United States Br. at 20-21 mentions three factors studied by Astin: (1) institutional diversity emphasis, including a commitment to increasing the number of minority faculty and students; (2) multiculturalism in the general education curriculum, and (3) direct student experience with diversity, including taking ethnic studies courses, attending cultural awareness workshops, socializing with other-race students and discussing racial issues with peers. The passage cited by Judge Duggan says nothing about racial diversity itself. The same is true for the passages cited by Duggan from the other two briefs.

Duggan noted in his opinion that the NAS specifically took issue with the studies relied upon by the American Association of Law Schools, contending that such studies really report that 'outcomes are generally not affected' by racial diversity on campus (NAS Br. at 6-7)."[66] *Judge Duggan never disputes this claim in his opinion.* This is quite inexplicable on the merits, since the University's defense collapses completely if outcomes are *not* affected by racial diversity on the campus. This lapse in the argument can only be explained if one assumes that Duggan took the question whether diversity *activities* are connected with positive educational outcomes to be the same thing as the question whether the racial diversity of the student body is. But clearly, they are not the same thing.

There can be no doubt that Judge Duggan correctly attributed to the University Defendants the claim that there are "educational benefits that flow from a racially and ethnically diverse student body"[67], for there is simply no other way of reading the following passage from Gurin's Expert Witness Report:

> A racially and ethnically diverse student body has far-ranging and significant benefits for all students, non-minorities and minorities alike. Students learn better in a diverse educational environment, and they are better prepared to become active participants in our pluralistic, democratic society once they leave such a setting.[68]

What Judge Duggan missed, however, is that Gurin fails to show anything of the sort. Indeed, in a response to the NAS' brief dated January 11, 2001 (i.e., after the Duggan Opinion was issued), Gurin clearly recedes from this claim:

> Structural diversity [her own term for racial diversity] is essential but, by itself, usually not sufficient to produce substantial benefits; in addition to being together on the same campus, students from diverse backgrounds must also learn about each other in the courses that they take and in informal interaction outside of the classroom. *For new learning to occur, institutions of higher education have to make appropriate use of structural diversity.* (italics in original).

Note that there is no assertion here that racial diversity is associated with positive educational outcomes (the claim on which Judge Duggan largely based his opinion). It turns out that what Gurin is claiming (and

[66] Duggan Opinion: 25.
[67] Ibid., p. 21.
[68] Gurin 1998: 3 (Summary and Conclusions).

in fact claimed all along)–despite the clear, unambiguous, and highly misleading statement from the Expert Report that Duggan was no doubt echoing in his opinion–is that structural diversity is correlated with positive educational outcomes, but only when it is conjoined with other factors that may or may not be present on campus.[69]

The really crucial point, however, and the one that completely devastates the University's defense of its racial discrimination, is simply this: *Gurin doesn't demonstrate the latter, revised proposition either*. In fact, the CIRP database tells us that one must reject the claim that campus racial diversity is connected with the alleged educational benefits even when structural diversity is conjoined with Gurin's four "campus experience variables."

If it were true, the claim that structural diversity produces educational benefits when conjoined with Gurin's four "campus experience variables" would actually be of some interest. For example, if the University could establish that racially diverse ethnic studies courses produce educational benefits and also that *only* racially diverse ethnic studies courses on racially diverse campuses do so, then the University would have established that campus racial diversity has an indirect positive effect on student outcomes, or to put this another way, that it has an effect, but only in the presence of one or more mediating variables. But the University cannot make this argument, because the four campus experience variables that Gurin considers (i.e., the things that Astin typically calls "diversity activities") are *controlled for in the ACE-HERI-CIRP regressions*, and because we know that after all these variables have been controlled for, the regressions fail to find either direct *or indirect* correlations between racial diversity and final student outcomes. This means that Gurin cannot argue that racial diversity produces educational benefits *even when it is conjoined with these other factors*.

In Gurin's scheme, the focus is, from first to last, on her campus experience variables rather than on the racial diversity of the institution. Whatever merits this may have from the point of view of educational theory, it is nonsense to place the emphasis where she does when it is a matter of testing the hypothesis that *campus racial diversity* is connected with educational benefits, which is the only issue that is before the courts.

In Step 3 of her regressions, Gurin tests whether the four campus experience variables she considers continue to have statistically significant effects after controlling for structural diversity (i.e., she tests

[69] In our view, this means that Gurin's unqualified statement, which she no longer holds (if she ever did), should be stricken from the court record.

whether or not the campus experience variables have an impact on educational outcomes no matter what the racial diversity of the student body is). This, however, is a matter that is irrelevant to the constitutional question before the courts. On other hand, she *fails* to test whether the racial diversity of the student body is correlated with student outcomes, even though this test is *directly* on point for the legal issues. Furthermore, she fails to do this even though the CIRP database permits this crucial test, and despite the fact that Astin in his full-scale report of the CIRP database for the same 1985-89 longitudinal study ran this test and found that the database tells us to *reject* the very claim that the University is making in court! Finally, as Astin found, there are no direct *or* indirect correlations of any educational significance to be found between the racial diversity of the student body and student outcomes. Thus, the very database that Gurin uses tells us to reject the hypothesis that lies at the core of her analysis. And although she must have been aware of these findings, she does not even so much as mention them in her 160-page "Expert Witness Report"–the very document that comprises the heart and soul of the University of Michigan's claims about the educational benefits of campus racial diversity.

CONCLUSION

The University of Michigan is ideologically and institutionally committed to racial preferences in admissions. It is entitled to advance what it sees as moral and legal arguments that support its position. But there is no justification for a misleading statistical analysis. There is no justification for twisting data that refute claims about the educational value of diversity in order to make it appear that these data support such claims.

BIBLIOGRAPHY

BIBLIOGRAPHY

ACE Brief. Brief of *Amici Curiae* American Council on Education, et al. in support of defendants' motion for summary judgment and in opposition to plaintiffs' motion for summary judgment, in Gratz, et al. v. Bollinger, et al., No. 97-75321 (E.D. Mich.).

Adam, Nabil R. and John C. Wortmann. 1989. Security-control methods for statistical databases: A comparative study. ACM Computing Surveys, Vol. 21, No. 4, December 1989: 515-556.

American Council on Education. 1953. A brief statement of the history and activities of the American Council on Education, 1918-53. July 1953. Washington, D.C.: American Council on Education.

Astin, A. W. 1993a. Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. Phoenix, AZ: American Council on Education, Oryx Press.

Astin, A. W. 1993b. Diversity and multiculturalism on campus: How are students affected? Change 23, 44-49.

Astin, A. W. 1993c. What matters in college?: Four critical years revisited. San Francisco: Jossey Bass. [The hardback edition was published in 1991, New York: American Council on Education / Macmillan Publishing Company]

Astin, Alexander W. 1977. Four critical years. San Francisco: Jossey-Bass.

Astin, Alexander W., Korn, William S., and Dey, Eric L. 1991a. The American college teacher: National norms for the 1989-90 HERI faculty survey. Los Angeles: Higher Education Research Institute, University of California.

Astin, Alexander W., Treviño, Jesús G. and Wingard, Tamara L. 1991b. The UCLA campus climate for diversity: findings from a campuswide survey conducted for the Chancellor's Council on Diversity. Los Angeles: Higher Education Research Institute, University of California.

Astin, A. W., Panos, R. J., and Creager, J. A. (1966). A program of longitudinal research on higher education. ACE Research Reports, 1(1). Washington, DC: American Council on Education. [ERIC Document Reproduction Service Order Number ED 030 362]

Bowen, W. G., & Bok, D. 1998. The shape of the river: Long-term consequences of considering race in college and university admissions. Princeton, NJ: Princeton University Press.

Chang, M. J. (manuscript, n.d.) Does increase [sic] racial diversity lead to a more diverse collection of thoughts, ideas, and opinions on campus?: A

study of racial diversity and students' viewpoints. http://www.gseis.ucla.edu/faculty/chang/Diversity_Viewpoints.pdf

Chang, M. J. 1999. Does racial diversity matter?: The educational impact of a racially diverse undergraduate population. Journal of College Student Development 40(4) 377-395.

Chang, M. J. 1996. Racial diversity in higher education: Does a racially mixed student population affect educational outcomes? Unpublished doctoral dissertation, University of California, Los Angeles. Ann Arbor, MI: University Microforms International (UMI), Order Number 9626812.

Chang, M. J., Witt, D., Jones, J., & Hakuta, K. (eds.) (in press). Compelling interests: Examining the evidence on racial dynamics in higher education. Palo Alto, CA: Stanford University Press. A report of the AERA panel on racial dynamics in colleges and universities. (Prepublication draft advance copy: http://www.stanford.edu/~hakuta/racial_dynamics/Compelling1.pdf)

Denning, Dorothy E. and Jan Schlörer. 1983. Inference controls for statistical databases. Computer 16, 7 (July), 69-82.

Duggan Opinion (December 13, 2000). The Hon. Patrick J. Duggan, United States District Judge, United States District Court, Eastern District of Michigan, Southern Division, Case No.: 97-CV-75231-DT.

Freedman, D. R. Pisani and R. Purves. 1998. Statistics. W.W. Norton & Co, Third Edition.

Gurin, P. (11 January 2001) "Supplemental Expert Report of Patricia Y. Gurin." *Grutter, et al. v. Bollinger,* et al. Case No. 97-75928.

Gurin Deposition. 1999. Deposition of Patricia Gurin, February 12, 1999 in the matter of: Barbara Grutter v. Lee Bollinger, et al. Esquire Deposition Services, Ann Arbor, MI. Original File 021299PG.ASC, 90 pages, Min-U-Script® File ID: 1373710982.

Gurin, P. 1998. Expert witness report of Patricia Y. Gurin, in Gratz, et al. v. Bollinger, et al., No. 97-75321 (E.D. Mich.). Ann Arbor: The University of Michigan. December 15, 1998. (Online version available on the Web at http://www.umich.edu/~urel/admissions/legal/expert/gurintoc.html)

Lynch, F. R. 1997. The diversity machine: The drive to change the 'white male' workplace. New York: The Free Press. (pp. 274-323, "Multicultural vision meets multiversity realities: The University of Michigan.")

The Michigan study project. Office of Multicultural Initiatives. Ann Arbor: University of Michigan. 1994.

National Association of Scholars. (14 July 2000). Brief of National Association of Scholars as amicus curiae in support of plaintiffs' motion for partial summary judgment. *Gratz, et al., v. Bollinger, et al.,* No. 97-75321 (E.D. Mich.). National Association of Scholars: Princeton, N.J.

Pascarella, E. T. and Terenzini, P. T. 1991. How college affects students: Findings and insights from twenty years of research. San Francisco: Jossey-Bass.

Pascarella, E. T., Edison, M., Nora, Al, Hagedorn, L. S., & Terenzini, P. T. 1996a. Influences on students' openness to diversity and challenge in the first year of college. Journal of Higher Education. 67(2) 174-195.

Pascarella, E. T., Whitt, E. J., Nora, A., Edison, M., Hagedorn, L. S., & Terenzini, P. T. 1996b What have we learned from the first year of the National Study of Student Learning? Journal of College Student Development 37(2): 182-192.

Sax, Astin, Korn, & Gilmartin. 1999. The American college teacher: national norms for the 1998-99 HERI faculty survey. Los Angeles: Higher Education Research Institute, University of California.

Tendick, Patrick and Norman Matloff. 1994. A modified random perturbation method for database security. ACM Transactions on Database Systems, Vol. 19, No. 1, March 1994: 47-63.

Whitt, E. J., Edison, M. I., Pascarella, E. T., Terenzini, P. T., and Nora, A. 1998. Influences on students' opennness to diversity and challenge in the second and third years of college. Paper presented at the Annual Meeting of the Association for the Study of Higher Education.

Wingard, T. L., Treviño, J. G., Dey, E. L., and Korn, W. S. 1991. The American college student, 1989: National norms for 1985 and 1987 college freshmen. Los Angeles: Higher Education Research Institute, University of California.

Wood, Thomas E. 1999. Good enough for the Supreme Court? Review of *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*, by William G. Bowen and Derek Bok. Academic Questions, Spring 1999, Vol. 12, No. 2, pp. 82-87. Rutgers, N.J.: Transaction Periodicals Consortium.

ABOUT THE AUTHORS

THOMAS E. WOOD

Thomas Wood has been Executive Director of the California Association of Scholars, a state affiliate of the National Association of Scholars, since 1991. He is also a principal of Americans Against Discrimination and Preferences, a 501(c)(3) organization that is dedicated to promoting the abolition of racial and gender discrimination and preferences at the local, state, and federal levels, along the lines established by California's Proposition 209.

Wood and his colleague, Glynn Custred, were the authors and official proponents of California's Prop. 209. They began work on the initiative in November, 1991. Since that time, Wood has spoken to a wide range of groups on affirmative action, diversity, and race and gender issues in America. He has appeared many times on CNN, National Public Radio, and all three major TV networks. He has also spoken on these issues on local radio, college campuses, and to a wide range of public interest organizations.

Wood is the author of "Does decisional law grant whites fewer political rights under the Fourteenth Amendment than it grants to racial minorities?: A response to Vikram D. Amar and Evan H. Caminker" (Hastings Constitutional Law Quarterly, Summer 1997); the co-author, with Victor A. Walsh, of a California Association of Scholars study, *Bakke and Beyond* (1996), the only comprehensive study of preferences in California public higher education in the pre-209 world; a 1993 study issued by the California Association of Scholars ("Accreditation and diversity: A critique of WASC's "Draft Report on Diversity"); and "Good enough for the Supreme Court? A review of *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*, by William G. Bowen and Derek Bok" Academic Questions, Spring 1999.

Wood received his B.A. from the University of California, Berkeley in 1969, where he was elected to Phi Beta Kappa. He received his Ph.D. in philosophy from the University of California, Berkeley in 1975, where he was the recipient of a Ford Foundation Fellowship. He is the author of three scholarly works that have been published by the University of Hawaii Press for the Society for Asian and Comparative Philosophy.

MALCOLM J. SHERMAN

Malcolm J. Sherman is an associate professor in the Department of Mathematics and Statistics at the University at Albany, one of the four university centers of the State University of New York. He holds a joint appointment in the Department of Biometry and Statistics at the University at Albany School of Public Health. He has written research papers in mathematics and statistics, and has also authored articles and reviews on a wide range of mathematical, statistical, and public policy subjects. His mathematical research has been supported by the National Science Foundation. He teaches mainly graduate and undergraduate courses in statistics, and directs the actuarial program at the University at Albany.

He has served as Assistant to the Staff Director of the United States Commission on Civil Rights, and has taught courses in and written articles on civil rights. He is currently president of the New York Association of Scholars.

He received his Ph.D. in mathematics from the University of California, Berkeley and is a Phi Beta Kappa graduate of the University of Chicago.