

Description of the UTD Texas Schools Project

The University of Texas at Dallas (UTD) Texas Schools Project is a multiyear research project whose goals are to obtain a better understanding of the determinants of student performance with the long-term objective of providing a knowledge/research base to improve the performance of public schools. In spring 1999, the project broadened its scope to include an extensive program of research on minority access to Texas public colleges and universities.

John F. Kain, Cecil and Ida Green Chair for the Study of Science and Society at UTD, initiated the UTD Texas Schools Project in 1992 when he was a Visiting Professor at UTD. Prior to accepting a permanent UTD appointment in spring 1997, Professor Kain was the Henry Lee Professor of Economics and Professor of Afro-American Studies at Harvard, where the project was previously housed. It is now housed at UTD's Cecil and Ida Green Center for the Study of Science and Society.

Phase I: Creation of the Texas Schools Microdata Panel (TSMP)

The Spencer Foundation, which has provided nearly \$400,000 of funding for the development of the Texas Schools Microdata Panel (TSMP), supported the collection of the first eight years of data, and the difficult and time consuming effort of creating TSMP from the Texas Education Agency's (TEA) disparate and unlinked annual data. As we discuss in greater detail under Phase II, Spencer also funded the project's first substantive focus, an investigation of the impact of increased minority access to suburban schools on the academic performance of minority, and especially Black, children.

The version of TSMP implemented during Phase I includes enrollment and attendance data, as well as standardized test scores, for more than two million students for up to 10 years, 1990-1999.¹ TSMP begins in the 1989-90 school year because TEA

¹ PEIMS is a yearly relational data base and TEA makes no effort to link these data across years. To construct TSMP, we had to combine annual PEIMS teacher and student data with TAAS, NAPT, and various teacher certification tests that are not part of PEIMS, and link these data across years. As Appendix Table A-1 reveals, to create the Phase I version of TSMP we had to combine data from more

implemented its PEIMS (Public Education Information Management System) in that year. In each subsequent year, TEA has improved the quality and extent of these data.

As Table 1 reveals, the students included in the Phase I version of TSMP belong to five cohorts. The members of the youngest cohort were in Pre-K during the 1989-90 school year while members of the oldest were in third grade in the same year. The letter and number designations in the columns labeled Test/Grade in Table 1 identify particular standardized tests by grade and type of test NAPT (Norm-referenced Assessment Program for Texas), TAAS (Texas Assessment of Academic Skills) and TEAMS (Texas Educational Assessment of Minimum Skills). Thus, N-5 under Cohort 1 refers to the fifth grade NAPT, while TA-7 under Cohort 1 refers to the eighth grade TAAS and TE-1 under Cohort 3 refers to the first grade TEAMS. This test is the only first grade statewide test during the period covered by the Phase I TSMP.

In addition to student data, TSMP includes individual data for all Texas public school teachers for the same nine-year period. For all grades and years we are able to link these teacher data to individual students at the campus, grade and program [bilingual, ESL (English as a Second Language), special education, gifted and talented] level. For the seven grades/years for which NAPT was given, we have been able to link individual students to their specific teachers, albeit with some error.

For all students included in TSMP, we have been able to complete educational histories for individual students for all the years they attended Texas public schools. Use of these data has enabled us to more accurately and effectively assess the performance of Texas schools than could be done with the fragmentary data that was previously available. These data have also enabled us to develop a better understanding of the causes of low student performance.

than 140 individual student files and more than 110 individual teacher files, as well as campus level data from TEA's AEIS files, block group data from the 1990 Census and district level data from the School District Data Book CD ROMs. The number of individual records included in the Phase I TSMP exceeds 102 million.

Table 1. Total Students and Standardized Tests Included in the Texas Schools Microdata Panel (TSMP) by Cohort, Grade and Test

Year	Total Students (Enrollment)	Sem	Cohort 1		Cohort 2		Cohort 3		Cohort 4		Cohort 5	
			Gr	Test/ Grade								
89-90		F	3		2		1	TE-1	K		PK	
89-90	1,161,358	S	3		2		1		K		PK	
90-91	1,505,551	F	4		3	T-3	2		1		K	
90-91	1,391,735	S	4		3		2		1		K	
91-92	1,420,295	F	5		4		3	T-3	2		1	
91-92		S	5	N-5	4	N-4	3	N-3	2		1	
92-93	1,415,593	F	6		5		4		3	T-3	2	
92-93		S	6	N-6	5	N-5	4	N-4 & T-4	3	N-3	2	
93-94	1,428,908	F	7		6		5		4		3	
93-94		S	7	T-7	6	T-6	5	T-5	4	T-4	3	T-3
94-95	1,438,632	F	8		7		6		5		4	
94-95		S	8	T-8	7	T-7	6	T-6	5	T-5	4	T-4
95-96	1,459,220	F	9		8		7		6		5	
95-96		S	9		8	T-8	7	T-7	6	T-6	5	T-5
96-97	1,493,782	F	10		9		8		7		6	
96-97		S	10	T-10	9		8	T-8	7	T-7	6	T-6

While the TEA data are of unprecedented quality and extent, important gaps remain. Therefore, we have made extensive efforts to obtain supplementary data from individual school districts. Thus far, our highest priorities have been to add early (first and second grade) tests for students already included in our data base, information that will enable us to link individual students to individual teachers, and additional family background variables. We have already held discussions with officials in 17 districts in the Austin, Corpus Christi, Dallas, Fort Worth, Houston, San Antonio and Waco PMSAs and with the Catholic Diocese of Dallas. Six of these districts have already provided us with supplementary data and two others have agreed to participate. We are continuing our discussions with the remaining districts and plan to meet with others as time allows.

The students currently included in TSMP are getting older. In spring 1999, the youngest students included in TSMP will be finishing the eighth grade and the oldest will be graduating from high school. As funding permits, we will add younger cohorts to TSMP. If we were to make these additions at the present time, they would total some 45 grade/years and more than two million additional students. TSMP already includes teacher data for all years and grades.

Phase II. Research on Minority Suburbanization and It's Impact on Student Achievement

As noted previously, grants from the Spencer Foundation funded both the creation of TSMP and a program of research on differences in mean achievement levels for Anglo (white non-Hispanic), Asian, African American, and Native American children attending Texas public schools and the causes of these differences. This ongoing research program has given particular attention to the extent of minority suburbanization and to how much black students have benefited in terms of higher scores on standardized tests from greater access to higher quality suburban schools. Although these are complex and methodologically difficult issues, still preliminary findings presented in two working papers by John F. Kain and Daniel M. O'Brien, indicate the impact is substantial. These papers and, with one exception, all other UTD Texas Schools Project Working papers can be viewed at the Green Center website (www.utdallas.edu/research/greenctr) and

downloaded at no cost. The exception, a paper by John F. Kain and Kraig Singleton, “Equality of Educational Opportunity Revisited,” was published in the May/June 1996 issue of the *New England Economic Review*.

O’Brien’s research on summer fallback is described in the TSP working paper, “Family and School Effects on the Cognitive Growth of Minority and Disadvantaged Elementary School Students” included at this website. His research demonstrates that both black and Hispanic children start school with significantly fewer skills than Anglo students, that they on average have larger gains than than Anglo students during the school year, but that these larger gains are dissipated during the summer vacations when they fall back. Consequently, the gaps between black and Anglos and Hispanics and Anglos stay relatively constant and may even widen somewhat. This research on the importance of school vacations provides an excellent example of the value of adding data from individual school districts to the extensive and highly valuable data obtained from state agencies.

With the exception of cohort 3, the earliest standardized scores included in TSMP are for the third grade. A great deal, however, occurs between the kindergarten and the third grade. O’Brien uses first grade scores for TEAMS and other first grade scores obtained from individual districts to examine the relationship between first and third grade scores and between first and seventh grade scores.

Kain & O’Brien’s working paper, “A Longitudinal Assessment of Reading Achievement: Evidence from the UTD Texas Schools Project,” uses TSMP data for one cohort of students to assess the reasons for the large differences in reading achievement for African Americans, Asian Americans, Hispanics and Anglos (non-Hispanic whites).

Phase III. Research on Special Education

In fall 1996 the Smith-Richardson Foundation made a \$200,000 grant to Eric A. Hanushek (Rochester University) to use TSMP for a program of collaborative research by Hanushek, Kain, and Steven G. Rivkin (Amherst College) on special education. Heretofore, special education, which is the most rapidly growing segment of public

education, has been subject to very little systematic research. The appendix lists two working papers that report the findings of this special education research. The Smith-Richardson grant also funded significant research on related issues. This includes research that quantifies the relative contributions of individual teachers and schools to individual academic achievement, research that attempts to answer the question as to whether higher salaries buy better teachers and research on the cost to individual students of switching schools. These working papers may be viewed at this site and downloaded at no cost.

Phase IV. Research on Minority Access to Texas Colleges and Universities

In March 1999, The Andrew W. Mellon Foundation provided the UTD Texas Schools Project with a \$1.25 million grant to support research on the factors affecting minority students' decisions to attend college and on the effects of the recent court ruling barring the use of race in college admissions. This study will extend research described in an earlier Mellon Foundation funded study by William Bowen and Derek Bok. *The Shape of the River: Long Term Consequences of Considering Race in College and University Admission* focused on the admission to and educational experiences of African Americans attending twenty-eight "selective" colleges and universities. This research adds Hispanics and greatly broadens the range of institutions studied to include all Texas public colleges and universities, including junior and community colleges. Our earlier and ongoing research on racial/ethnic differences in performance on standardized tests in elementary, middle and high schools, should help us identify the causes of low high school grades and low SAT and ACT scores. "Selective" colleges and universities as well as less selective colleges and universities make extensive use of both grades and scores on college admission tests in making admission and financial aid decisions.

Our original plan for TSMP envisioned following the original five cohorts into college, for those who attended college, and then into the labor force. For those that did not enroll in higher education, we hoped to obtain labor force information as soon as they left school, if not before. The Mellon grant has enabled us to accelerate our earlier research plan by providing funding to add TEA data for nine cohorts of older students to TSMP, by obtaining Higher Education Coordinating Board data for all individuals enrolled in Texas public colleges and universities between 1990 and the present. We have also

begun negotiations with the Texas Workforce Commission about obtaining employment and earnings data for all individuals included in an expanded TSMP.

Of the cohorts of TEA data being added to TSMP, the youngest were in fourth grade in 1989-90 and the oldest were in 12th grade in 1989-90. These data acquisitions will add 57 years/grades of individual student data to the 45 grades/years of data already

In spring 1998 we created a supplemental data base to analyze ethnic and racial differences in graduation, dropout rates and course completions for students attending public high schools (Kain, 1998b). This modest study, which we completed for the Texas Commission on a Representative Student Body (the Hopwood Commission), is a precursor to the far more extensive analyses we will complete under the Mellon grant. Our research for the Hopwood Commission used course completion data, a listing of all courses taken and successfully completed. We anticipate we will make extensive use of these data in this study in an effort to document the academic programs of minority high school students and the effects of their choices on access to and success in college. We also anticipate that the availability of college preparatory and particularly advanced placement courses will vary by type of high school.

Because of its large size and ethnic and racial diversity, Texas is an incredible laboratory to study the way in which educational opportunities and choices affect educational outcomes for minority and low-income children and affect their access to and success in higher education. We will begin by quantifying the enrollment patterns by race/ethnicity and income and their changes over time. Then, using data on the pre-college educational histories of students included in TSMP, we will document and describe the ways in which constraints on residential and other choices of minorities tend to reduce their academic achievement in elementary, middle and high school and thus limit their opportunity to enroll in college.

More than 934,000 students were enrolled in institutions of higher education in Texas in fall 1997. Of those who were enrolled in public institutions, slightly more attended community/junior colleges than attended Senior Colleges and Universities. In addition, only eight percent of Texas residents who were enrolled as freshmen in some institution of higher learning in fall 1994 attended an out-of-state institution (US Department of Education, 1997).

The Coordinating Board maintains extensive data on all students enrolled in Texas public institutions of higher learning. The student IDs used by the Coordinating Board can be used to link these data to the individuals we have already included in TSMP. We have obtained these data for all students attending Texas public universities, senior

colleges and community/junior colleges from 1990 to the present. For some of these students, we will be able to construct complete educational histories from the time they enter kindergarten until they graduate from college.

The Coordinating Board does not obtain individual data for private colleges and universities, but it does collect these data for all public institutions of higher education and nine out of 10 Texas students participating in higher education are enrolled in public institutions.

The Coordinating Board data will include enrollment in particular programs and courses and individual student scores on TASP (Texas Academic Skills Program). To be accepted in a degree program at any Texas public educational institution, all applicants must take and pass TASP unless they have obtained sufficiently high scores on the ACT, SAT or TAAS writing, math and reading tests. To obtain an ACT or SAT exemption from TASP, applicants must have a composite ACT score of 23 or better (with a minimum of 19 on both the English and math tests) or a combined SAT score of 1070 (with a minimum of 500 on both the verbal and math tests).

The sorting by race/ethnicity and income of Texas high school graduates among Senior Colleges/Universities and community colleges will be an important focus of the proposed research. Similarly, the extent to which community colleges are effective pathways to four-year institutions and to success in the labor market for minority group members and for students from low-income households will be a core issue.

Support of Ph.D. Dissertations

Two UTD graduate students have received financial support from the Green Center and The UTD Texas Schools Project. Their research combines TSMP data with supplementary data obtained from one or more school districts.

Daniel M. O'Brien completed his dissertation in summer 1999 and is currently Associate Director of the UTD Texas Schools Project. His dissertation research focused on three areas. They are the effects of summer fallback on the achievement of low income and minority students, the effects of early tests on student achievement and the use of

locally collected data obtained from individual districts to augment the individual and family background variables that are currently being collected by TEA. Summer fallback refers to the summer achievement declines that appear to occur for low-income children during vacation periods when middle and high-income children continue to experience gains. O'Brien's research on the effects of early tests and the development of more extensive individual and family measures are useful in assessing the biases that may arise in the UTD Texas School Project research from the lack of these data in the larger TSMP database.

Sharon Wrobel is studying bilingual education in one of the states largest districts and will help that district assess its efforts to strengthen its bilingual programs. Like O'Brien, Wrobel has been able to supplement TSMP data with data obtained from the district she is studying. These data include the scores obtained by Limited English Proficient (LEP) students on the English proficiency tests that are used in determining whether they should be assigned to bilingual, ESL or regular programs. A test of this kind is given to all LEP students attending public schools, but the scores are not supplied to TEA. Thus, they are not included in TSMP.

Pro-Bono Research and Evaluation Activities

In addition to the data sharing arrangements with individual districts described above, the UTD Texas Schools Project is involved in a growing number of pro-bono efforts aimed at helping individual districts and other organizations assess programs and policies.

The best example of our pro-bono assistance to individual districts is our ongoing relationship with Fort Worth ISD (FWISD). During the past year, O'Brien and Kain helped FWISD design and carry out an evaluation of direct instruction reading programs they have implemented in 32 low achieving schools. FWISD is using Open Court in 14 of these schools and Reading Mastery in 18 others. The district wants to answer two questions. Do these programs provide measurable improvements in the reading skills of FWISD students? If both improve the reading skills of FWISD students, which is better?

FWISD has a talented and hard working assessment/evaluation staff, but they are somewhat lacking in the statistical methods that are needed to complete quantitative program assessments. These quantitative skills, which are our stock in trade, are thus a valuable complement to those of FWISD's assessment and evaluation staff. Sharon Wrobel's research on bilingual education is also designed to assist FWISD in their efforts to evaluate recent changes in the bilingual education programs. Limits on staff time and other resources have thus far limited our efforts to assist other districts in this way. We would like to expand this feature of our program, however, and are currently seeking additional funding that would permit us to expand these activities.

In spring 1999, Dr. Hardy Murphy, formerly Associate Superintendent at FWISD and a member of the Hopwood Commission and currently Superintendent of the Evanston-Skokie Schools (K-12), asked us to prepare an analysis of the experiences of high school students for the Commission. To complete these analyses we obtained supplementary data for two cohorts of high school students from TEA and used them to complete these analyses. The cost of acquiring additional data from TEA was borne by the UTD central administration. The resulting analyses are contained in a Texas Schools Project working paper "Ethnic and Racial Differences in Graduation, Dropout Rates and Course Completions for Students Attending Texas Public High Schools," that may be viewed at, or downloaded, the Green Center website.

"Teach for America," which, to our great surprise, has placed 528 corps members in Houston and 217 in the Valley, has asked us to use TSMP data to compare the retention/attrition rates of for their teachers to other first and second year teachers employed in Houston and in the Valley. We have agreed.

Our growing experience with individual districts and the recent request from "Teach for America" to help them evaluate their program in Texas strongly demonstrates the synergy between our successful development of TSMP, our program of research based on TSMP and collaborations with individual districts. With more than one thousand districts, Texas is an incredible laboratory to study educational policy and to identify programs and policies that work.

Our work for FWISD on bilingual education and reading, the report for the Hopwood Commission and the analyses we expect to complete for “Teach For America” were all completed at no cost to them. Green Center and Spencer Foundation funds covered these costs.

Publications and Working Papers

John F. Kain and Kraig Singleton. “Equality of Educational Opportunity Revisited.” *New England Economic Review*. (May/June), 1996.

John F. Kain and Daniel M. O’Brien, “Has Moving to the Suburbs Increased African American Educational Opportunities?” January 23, 1998.

John F. Kain and Daniel M. O’Brien, “A Longitudinal Assessment of Reading Achievement: Evidence from the UTD Texas Schools Project,” April 2-4, 1998.

John F. Kain, “Using TEA Annual Data to Develop a Multi-Year Panel Data Base: Lessons Learned and Suggested Additions and Improvements to TEA’s Data Collection,” May 11, 1998.

John F. Kain, “Ethnic and Racial Differences in Graduation, Dropout Rates and Course Completions for Students Attending Texas Public High Schools,” June 4, 1998.

Steven G. Rivkin, Eric A. Hanushek and John F. Kain, “Teachers, Schools and Academic Achievement,” August 1998.

Eric A. Hanushek, John F. Kain and Steven G. Rivkin, “Does Special Education Raise Academic Achievement for Students with Disabilities,” August 1998.

Eric A. Hanushek, John F. Kain and Steven G. Rivkin, “Does Special Education Work,” September 1998

Daniel M. O’Brien, “Family and School Effects on the Cognitive Growth of Minority and Disadvantaged Elementary School Students,” November 8-10, 1998.

John F. Kain, “The Impact of Individual Teachers and Peers on Individual Student Achievement,” October 29-31, 1998.

Sharon Leigh Wrobel and Daniel M. O’Brien, “Assessment of Bilingual Education Programs in a Large Texas School District,” October 31, 1998.

Eric A. Hanushek, John F. Kain and Steven G. Rivkin, “Do Higher Salaries Buy Better Teachers?” February 8, 1999.

Eric A. Hanushek, John F. Kain and Steven G. Rivkin, "The Cost of Switching Schools," April 1999.

Eric A. Hanushek, John F. Kain, Jacob M. Markman, and Steven G. Rivkin, "Do Peers Affect Student Achievement?" January 15, 2000.

John F. Kain and Daniel M. O'Brien, "Black Suburbanization in Texas Metropolitan Areas and Its Impact on Student Achievement," March 9, 2000.

February 9, 2000

Table A-1. Data and Files Included in the Texas Schools Microdata Panel (TSMP)
(Data for the 1990-97 School Years, Eight Years of Data)

File Types	Years	Files	Total Records
Student			
PEIMS Demographic	5	5	7,948,609
PEIMS Enrollment	8	8	11,147,832
PEIMS Chapter I Enrollment	8	8	5,330,209
PEIMS Special Ed Enrollment	8	8	1,465,578
PEIMS Voced Enrollment	4	4	986,627
PEIMS Gifted Enrollment	6	6	392,462
PEIMS Summer Demographic	5	5	7,948,609
PEIMS Basic Attendance	4	24	33,017,628
PEIMS Special Ed Attendance	4	24	6,984,950
PEIMS Voced Attendance	4	24	986,627
TAAS	7	22	6,259,435
NAPT	2	7	1,683,009
TEAMS	1	1	286,982
Total Student Files/Records			84,438,557
Teacher			
PEIMS Staff	8	8	2,421,138
PEIMS Employment	8	8	2,421,138
PEIMS Payroll	8	8	4,013,119
PEIMS Class	8	8	7,788,629
PEIMS Nonclass	8	8	366,579
PEIMS Permit	8	8	82,021
TECAT	5	1	145,711
ExCET	5	64	292,696
TOPT	5	1	4,894
PPST	5	1	54,125
TASP	5	1	32,032
Total Teacher Files/Records		116	17,622,082