

TEXAS HIGHER EDUCATION OPPORTUNITY PROJECT

Administrative College Application Data

Documentation for Public Use Data Files

December 18, 2008

CONTENTS

OVERVIEW	1
Introduction	1
Confidentiality.....	2
Variable Availability and Comparability	2
Admission Decisions.....	2
Variable Naming Convention	2
Missing Values.....	3
VARIABLE DESCRIPTIONS.....	5
Application Identifiers.....	5
Demographic Characteristics.....	6
Academic Characteristics	8
High School Characteristics.....	10
Admission and Enrollment Decisions	13
Graduation.....	15
APPENDIX A – SAT SCORE RECENTERING	16
College Board 1995-1996 SAT Score Recentering.....	16
SAT Variables and Recentering for THEOP Application Data	16
APPENDIX B – OVERLAPPING VALUES: RECODED TEST SCORE AND GRADUATION VARIABLES	17
Test Scores.....	17
Graduation Year	20
APPENDIX C – FIELDS OF MAJOR AND DEPARTMENTS OF MAJOR BY FIELD	21

OVERVIEW

INTRODUCTION

The Texas Higher Education Opportunity Project (THEOP) Administrative College Application Public Use Data (henceforth Application Data) consists of freshman application records for seven public and two private Texas universities. Seven of the nine universities provide all applicant records and two provide applicant records for enrollees only in specific years, as detailed in Table 1.

For the public institutions, Application Data coverage begins several years prior to the implementation of the Texas top 10% law (H.B. 588) in 1998, and extends until at least 2002. Application Data for the two private institutions is available for the period after implementation of the automatic admission law.

Table 1. Application Data: Participating Universities, Type, Years, and Size

University	Type	Application Data	
		Years	Number of Applications
Texas A&M	Public	1992-2002	163,027
Texas A&M Kingsville*	Public	1992-2002	18,872
UT Arlington	Public	1994-2002	29,844
UT Austin	Public	1991-2003	210,006
UT Pan American	Public	1995-2002	44,747
UT San Antonio**	Public	1990-2004	61,221
Texas Tech	Public	1995-2003	81,153
Rice	Private	2000-2004	36,190
Southern Methodist	Private	1998-2005	45,549

* Application data for enrollees only: 1992-1994

** Application data for enrollees only: 1990-1997

Each university was asked to provide a subset of the items typically found on a college admission application. These items included the year and term an applicant desired to enroll, applicant demographics (gender, race, citizenship, Texas residency), applicant academic characteristics (high school class rank, high school grade point average, SAT score, ACT score, AP classes taken, TOEFL score, desired major), and high school characteristics (high school name, address, city, state, senior class size). The universities were also asked to indicate admission and enrollment decisions, as well as college graduation dates for those enrollees who earned a degree within the time span of the data set. There is considerable overlap in information provided across institutions, but not every university provided identical sets of data items. Universities were not asked to provide applicant essays or specific high school course-taking information.

To increase the usefulness of the Application Data for research purposes, several data items are added to each application record. These include flags indicating whether an applicant's high school is (1) public or private, (2) regular or special, and (3) in Texas or out of state. Additionally, for

Texas high schools, the data include a variable indicating the percentage of economically disadvantaged students.

CONFIDENTIALITY

Our goal in preparing the Application Data is to make available the highest quality information while minimizing the risk of applicants being identified. Several confidentiality measures are taken to achieve this goal, including:

- *Elimination of identifiers*: University-assigned applicant identification numbers are changed to new, randomly generated applicant identification numbers. High school identifiers such as high school name, city and state are removed from the public use data files. Under special circumstances, researchers may apply for access to restricted use data through the THEOP website (<http://theop.princeton.edu>).
- *Collapsing small frequency cells*: Small frequency (less than 20) cells are eliminated by collapsing multiple values into range categories. For example, individual test score values are collapsed into test score ranges.
- *Setting values to missing*: Some values are set to missing to preserve the well-established categories but to hide individual values. This strategy is used infrequently to minimize its impacts on analyses.

VARIABLE AVAILABILITY AND COMPARABILITY

Table 2 shows variable availability for each university. Initials in a cell indicate that the variable is available for the university data file. Some variables are recoded to make them comparable across universities, but other variables cannot be rendered comparable and thus are available only for a subset of universities.

ADMISSION DECISIONS

The 1997 judicial ban on consideration of race and ethnicity in admission decisions applied to all public and private institutions, but the top 10% law, which *guaranteed* admission to students who graduated in the top 10% of their high school class, only applied to public institutions. All students were required to submit test scores for an application to be considered complete; however, standardized test scores were disregarded for rank-eligible applicants.

The Application Data shows that small numbers of in-state, top decile applicants were not admitted after the automatic admission law went into effect (1998). Because qualified applicants were guaranteed admission ***provided they submitted a completed application***, top 10% applicants not admitted were missing at least one or more requirements, such as test scores, application fees, essays or high school transcripts. Application fees, essays, and high school transcripts are not included in the application data.

VARIABLE NAMING CONVENTION

Some variables have a name that ends with a capital letter, R. This indicates that the variable has been recoded rather than appearing exactly as provided by the universities. Examples include SAT score (satR) and high school class rank in deciles (decileR).

MISSING VALUES

Items with a very large share of missing values are excised from all files. Examples of this include data items pertaining to AP test taking, TOEFL test scores and high school grade point average. Further, because UT Austin and Texas Tech data for early years contain many variables with a high percentage of missing values, the Application Data does not include applications for UT Austin for 1990 or Texas Tech for 1991-1994.

Table 2. Application Data: Variable Availability

Variable Name	Variable Label	Institution / Year								
		Texas A&M 1992- 2002	Texas A&M Kingsville 1992- 2002	UT Arlington 1994- 2002	UT Austin 1991- 2003	UT Pan American 1995- 2002	UT San Antonio 1990- 2004	Texas Tech 1995- 2003	Rice 2000- 2004	Southern Methodist 1998- 2005
studentid	Student id	am	amk	ar	au	pa	sa	tt	ri	smu
yeardes	Year admission desired	am	amk	ar	au	pa	sa	tt	ri	smu
termdes	Term admission desired	am	amk	ar	au	pa	sa	tt	ri	smu
termapp	Term of application	-	amk	-	-	-	-	-	-	-
male	Male	am	amk	ar	au	pa	sa	tt	ri	smu
ethnic	Ethnicity	am	amk	ar	au	pa	sa	tt	ri	smu
citizenship	Citizenship	am	amk	-	au	pa	sa	tt	-	smu
restype	Resident type	am	amk	ar	au	pa	sa	tt	-	smu
satR*	SAT composite score (Recentered)	am	amk	ar	au	pa	sa	tt	ri	smu
sat_not_recenteredR*	SAT composite score (Not recentered)	-	amk	ar	-	-	sa	tt	-	-
actR	ACT composite score	am	amk	ar	au	pa	sa	tt	-	smu
testscoreR*	SAT if provided, or ACT set to SAT Scale	am	amk	ar	au	pa	sa	tt	ri	smu
decileR	Student HS class rank: decile	am	amk	ar	au	pa	sa	tt	ri	smu
quartile	Student HS class rank: quartile	am	-	ar	-	-	-	-	-	-
major_field	Field of desired first choice major	am	amk	ar	au	pa	sa	tt	ri	smu
hsprivate	Private HS	am	amk	ar	au	pa	sa	tt	ri	smu
hstypeR	HS type: regular or other	am	amk	ar	au	pa	sa	tt	ri	smu
hsinststate	Texas HS	am	amk	ar	au	pa	sa	tt	ri	smu
hseconstatus	HS economic status	am	amk	ar	au	pa	sa	tt	ri	smu
hslos	HS in Longhorn Scholarship program	am	-	-	au	-	-	-	-	-
hscentury	HS in Century Scholarship program	am	-	-	au	-	-	-	-	-
admit	Admitted	am	amk	ar	au	pa	sa	tt	ri	smu
admit_ut_summer	Admitted: 3-way variable	-	-	-	au	-	-	-	-	-
admit_prov	Provisional admission	am	-	-	-	-	sa	tt	-	smu
utsa_cap	Coordinated admissions program (CAP)	-	-	-	-	-	sa	-	-	-
enroll	Enrolled	am	amk	ar	au	pa	sa	tt	ri	smu
gradyear	Graduation year	am	amk	ar	au	pa	sa	tt	ri	-

Note: Initials in the cell indicate that the variable is available for the institution.

* For these variables, availability by year differs by institution. See Appendix A for full description.

VARIABLE DESCRIPTIONS

APPLICATION IDENTIFIERS

This group of variables includes studentid, yeardes, termdes, and termapp.

Variable Name	Variable Label
studentid	Student id
Description Randomly generated identifier for each application.	
Notes Most institutions use a unique identifier for each application. Texas A&M, UT Arlington, and Southern Methodist use a single identifier for all applications submitted by a single student (for a different year and/or term). Student identifiers are not unique across institutions. Therefore, if building a dataset with application records from more than one institution, creation of an institutionid variable is recommended. This variable is required to merge Application Data with associated College Transcript data.	

Variable Name	Variable Label
yeardes	Year admission desired
Description Calendar year of desired admission.	

Variable Name	Variable Label
termdes	Term admission desired
Description Academic term of desired admission.	
Values	Value Labels
1	Spring
2	Summer I
3	Summer II
4	Fall
.	Missing
Notes All nine universities receive freshman applications for the fall term. Some universities also receive applications for Spring, Summer I or Summer II terms.	

Variable Name	Variable Label
termapp	Term application received
Description Academic term application is received.	
Values	Value Labels
1	Spring
2	Summer I
3	Summer II
4	Fall
.	Missing
Notes Because the termdes variable is missing for many applications in the Texas A&M Kingsville file, the termapp variable is also provided for this institution.	
Availability Available only for Texas A&M Kingsville.	

DEMOGRAPHIC CHARACTERISTICS

This group of variables includes male, ethnic, citizenship, and restype.

Variable Name	Variable Label
male	Male
Description Applicant sex.	
Values	Value Labels
0	Female
1	Male
.	Missing

Variable Name	Variable Label
ethnic	Ethnicity
Description Applicant race/ethnicity.	
Values	Value Labels
1	Black, Non-Hispanic
2	Hispanic
3	American Indian/Alaskan Native
4	Asian or Pacific Islander
5	White, Non-Hispanic
6	International
7	Other
.	Missing

Notes

Not all institutions use identical categories to report race and ethnicity. For example, Texas A&M Kingsville uses the categories Black, Hispanic, White, Other, while UT-Austin uses the categories Black, Hispanic, American Indian/Alaskan Native, Asian or Pacific Islander, White, International, Other. See institution-specific tabulations for details.

In addition to representing students who did not identify with one of the choices provided, Other includes students in categories that required collapsing in order to maintain applicant confidentiality. In cases where the number of International applicants is small, International is collapsed into Other.

The category International is used by institutions to identify students who are applying from non-US high schools. Note that applicants with an International value for the ethnic variable may or may not be US citizens. Missing is used for applicants who do not provide an indication of race/ethnicity.

variable Name	Variable Label
citizenship	Citizenship

Description
Applicant US citizenship status.

Notes
Texas A&M, Texas A&M Kingsville, Texas Tech, UT San Antonio, UT Pan American, and Southern Methodist provide a two-way classification:

Values	Value Labels
1	US Citizen
2	Non US Citizen

UT Austin provides a three-way classification:

Values	Value Labels
1	US Citizen
2	Non US Citizen
3	Permanent Resident Alien

Availability
Not available for UT Arlington and Rice.

Variable Name	Variable Label
restype	Resident type

Description
Applicant state resident type.

Notes
Texas A&M, Texas A&M Kingsville, Texas Tech, UT San Antonio, UT Pan American, and Southern Methodist provide a two-way classification:

Values	Value Labels
1	Texas Resident
2	Non Texas Resident

UT Austin and UT Arlington provide a three-way classification:

Values	Value Labels
1	Texas Resident
2	Non Texas Resident
3	Non US Resident

For institutions that provide only two categories, Non Texas Resident includes those who reside outside the US.

Availability
Not available for Rice.

ACADEMIC CHARACTERISTICS

This group of variables includes satR, sat_not_recenteredR, actR, testscoreR, major_field, decileR, and quartile.

Variable Name	Variable Label
satR	SAT composite score (Recentered)
Description Applicant SAT composite score.	
Notes SAT composite scores are collapsed to maintain confidentiality. See “Appendix B – Overlapping Values: Recoded Test Score and Graduation Variables” for a complete table of satR categories and values.	

Variable Name	Variable Label
sat_not_recenteredR	SAT composite score (Not recentered)
Description Applicant SAT composite score for early years that have not been recentered.	
Notes See “Appendix A – SAT Recentering” for a complete discussion.	
Availability Available only for Texas A&M Kingsville, UT Arlington, UT San Antonio, and Texas Tech.	

Variable Name	Variable Label
actR	ACT composite score.
Description Applicant ACT composite score	
Notes ACT composite scores are collapsed at the tails (e.g. “Less than 16” or “29-36”), with ranges of one or two in between (e.g. “21-21”). See “Appendix B – Overlapping Values: Recoded Test Score and Graduation Variables for a complete table of actR categories and values.	
Availability Not available for Rice.	

variable Name	Variable Label
testscoreR	SAT if provided, otherwise ACT set to SAT scale
Description	
Applicant SAT composite score, or if not available, applicant ACT composite score converted to equivalent recentered SAT composite score.	
Notes	
See “Appendix B – Overlapping Values: Recoded Test Score and Graduation Variables” for a complete table of testscoreR categories and values. Note that testscoreR is available for 1996 and later for most schools. testscoreR is available for 1997 and later for UT San Antonio, 1995 and later for UT Arlington and UT Pan American	
College Board ACT-SAT concordance table may be found at: http://professionals.collegeboard.com/research/pdf/rr9901_3913.pdf . (Table A.1; pg. 17).	

Variable Name	Variable Label
major_field	Field of desired first choice major (or Major for which student is being considered – UT Austin only)
Description	
Applicant’s intended field of major at time of application.	
Values	Value Labels
1	Agriculture
2	Architecture
3	Business
4	Education
5	Engineering/Computer Science
6	Fine Arts
7	General Studies
8	Health
9	Humanities
10	Individualized/Interdisciplinary
12	Natural/Physical Sciences
13	Other
14	Social Sciences
15	Social Work
16	Technical/Vocational
17	Undeclared
.	Missing

Notes

major_field specifies the field of intended major at time of application, not an enrollee’s actual major (which can be found in College Transcript Data). For UT Austin only, students apply and are accepted to a particular school (such as the Engineering School), so the variable description for UT Austin is “Field of Major for which student is being considered.” In infrequent cases, small cell sizes for certain fields in certain years are set to missing to maintain confidentiality. A complete list of majors within each field is available in “Appendix C – Fields of Major and Departments of Major by Field.”

Variable Name	Variable Label
decileR	Student HS class rank: decile

Description
Decile of applicant high school class rank.

Values	Value Labels
1	Top 10%
2	Second Decile
3	Third Decile
4	Fourth Decile
5	Fifth Decile
6	Sixth Decile
7	Seventh Decile and Below

Notes
For all schools except UT Arlington, decileR is calculated by dividing high school rank by high school senior class size. For UT Arlington, decileR is calculated from high school class rank percentage scores provided directly by the university. decileR is available beginning in 1999 for UT Arlington.

Variable Name	Variable Label
quartile	Student HS class rank: quartile

Description
Quartile of applicant high school class rank.

Values	Value Labels
1	Top 25%
2	Second Quartile
3	Third Quartile
4	Fourth Quartile

Availability
Available only for Texas A&M and UT Arlington.

HIGH SCHOOL CHARACTERISTICS

This group of variables includes hsprivate, hstypeR, hsinstate, hseconstatus, hslos, and hscenury.

Variable Name	Variable Label
hsprivate	Private HS

Description
Public/private designation of applicant high school.

Values	Value Labels
0	Public
1	Private
.	Missing

Variable Name	Variable Label
hstypeR	HS type: regular or other
Description	
Type of applicant high school, as defined by National Center for Educational Statistics (NCES).	

Values	Value Labels
0	Regular
1	Other
.	Missing

Notes
Other includes schools defined by NCES as Vocational, Special Education, and Alternative, as well as high schools that fall outside of these categories or cannot be clearly identified.

Variable Name	Variable Label
hsinstat	Texas HS
Description	
In-state (Texas) designation of applicant high school.	

Values	Value Labels
0	No
1	Yes
.	Missing

Variable Name	Variable Label
hseconstatus	HS economic status (Year specific)

Description
Economic status quartile of applicant high school. Available for regular, Texas public high schools 1994-2004.

Values	Value Labels
1	Upper quartile
2	Upper-middle quartile
3	Lower-middle quartile
4	Lower quartile
.m	Missing: TX Public HS
.n	Not available: not TX public HS 1994-2004

Notes
Publicly available data from the Texas Education Agency (TEA) is used to stratify regular, Texas public high schools according to the socioeconomic status of the students they serve.

The 25% of high schools with the lowest percent of students ever economically disadvantaged are designated as Upper quartile. The 25% of high schools having the highest percent of students ever economically disadvantaged are designated as Lower quartile. Because the statewide share of economically disadvantaged students rose over time, quartile cut points are calculated separately for each year. Table 3 reports the time-varying cut points that define each quartile.

Table 3. Percent Economically Disadvantaged Cut Points by Application Year

Application Year	25 th Percentile	50 th Percentile	75 th Percentile
1994	17.8	27.8	41.4
1995	18.6	29.0	43.0
1996	19.4	29.4	43.7
1997	19.5	29.5	43.9
1998	20.0	30.7	45.9
1999	20.0	31.6	47.1
2000	20.8	32.1	47.0
2001	20.6	32.4	47.7
2002	20.4	32.9	49.0
2003	22.2	34.6	51.0
2004	23.8	36.5	53.1

Variable Name	Variable Label
hslos	HS in Longhorn Opportunity Scholarship program (Year Specific)

Description
Applicant high school included in Longhorn Opportunity Scholarship program.

Values	Value Labels
0	No
1	Yes
.	Missing

Notes
This program began in 1999 and is sponsored by UT Austin. Included high schools have a low college-going tradition and typically have a significant percent of students from economically underprivileged backgrounds. Tuition scholarships are provided for a select number of students from these high schools who enroll at UT Austin. List of included high schools changes slightly each year. More information on the Longhorn Opportunity Scholarship program may be found at: <http://www.texascholarships.org/types/osfs/los.html>

Availability
Available only for UT Austin and Texas A&M.

Variable Name	Variable Label
hscentury	HS in Century Scholarship program (Year Specific)

Description
Applicant high school included in Century Scholarship program.

Values	Value Labels
0	No
1	Yes
.	Missing

Notes
This scholarship program began in 2000 and is sponsored by Texas A&M. Included high schools have a low college-going tradition and typically have a significant percent of

students from economically underprivileged backgrounds. Tuition scholarships are provided for a select number of students from these high schools who enroll at Texas A&M. List of included high schools changes slightly each year. More information on the Century Scholarship program may be found at: <http://centuryscholars.tamu.edu/CenturyScholar.html>

Availability

Available only for UT Austin and Texas A&M.

ADMISSION AND ENROLLMENT DECISIONS

This group of variables includes admit, admit_ut_summer, admit_prov, utsa_cap, and enroll.

Variable Name	Variable Label
admit	Admitted

Description
Institution's admission decision.

Values	Value Labels
0	No
1	Yes
.	Missing

Notes

The Application Data shows that small numbers of in-state, top decile applicants were not admitted after the automatic admission law went into effect (1998). Because qualified applicants were guaranteed admission ***provided they submitted a completed application***, top 10% applicants not admitted were missing at least one or more requirements, such as test scores, application fees, essays or high school transcripts. Application fees, essays, and high school transcripts are not included in the application data.

For UT Austin, when a student was denied admission, an offer was sometimes made to participate in a summer program that allowed fall enrollment upon successful completion. The variable admit_ut_summer indicates which applicants were given this opportunity.

Variable Name	Variable Label
admit_ut_summer	Admitted (3-way variable)

Description
UT Austin's admission decision indicating offer to participate in a summer program at UT Austin that allowed fall enrollment upon successful completion.

Values	Value Labels
0	No
1	Yes
2	Summer admission program
.	Missing

Notes

The UT Austin two-way admission variable (admit) shows that applicants given this summer opportunity are considered to have an admission decision of No.

Availability

Available only for UT Austin.

Variable Name	Variable Label
admit_prov	Provisional admission

Description

Institution's admission decision is provisional.

Values	Value Labels
0	No
1	Yes
.	Missing

Availability

Available only for Texas A&M, Texas Tech, UT San Antonio, and Southern Methodist.

Variable Name	Variable Label
utsa_cap	Coordinated admissions program (CAP)

Description

Applicant is a participant in Coordinated Admissions Program (CAP).

Values	Value Labels
0	No
1	Yes – CAP Applicant

Notes

Indicates freshmen who applied to UT Austin and were denied regular admission but were offered the opportunity to participate in the Coordinated Admissions Program (CAP). This program allows students to enroll at another UT system institution for their freshman year and if they complete 30 hours of coursework and earn a 3.2 GPA, they are allowed to transfer to UT Austin for their sophomore year. This program began in 2002.

Availability

Available only for UT San Antonio.

Variable Name	Variable Label
enroll	Enrolled

Description

Applicant's decision to enroll.

Values	Value Labels
0	No
1	Yes
.	Missing

GRADUATION

This variable indicates the year of graduation.

Variable Name	Variable Label
gradyear	Graduation year
Description	
Calendar year of college graduation.	
Values	
Value Labels	
1	1992, 1993, or 1994
2	1994 or 1995
3	1996 or 1997
4	1996, 1997, or 1998
5	1994
6	1995
7	1996
8	1997
9	1998
10	1999
11	2000
12	2001
13	2002
14	2003
15	2004
16	2005
17	2006
18	2007
21	Graduated by 2002-actual date not shown
22	Graduated by 2004-actual date not shown
23	Graduated by 2005-actual date not shown
24	Graduated by 2007-actual date not shown
.c	Censored
.n	Did not enroll

Notes

For Texas A&M, Texas A&M Kingsville, UT San Antonio, and UT Pan American, the earliest two or three graduation years are collapsed into one set of years to maintain confidentiality. To avoid very small cell sizes, in limited cases gradyear variable indicates “Graduated by [last year follow-up data provided]- actual year not shown.” The cases in this category are typically students who graduated in up to 2 years after admission or more than 6 years after admission. For enrolled students with unknown graduation status due to finite follow-up period, gradyear variable indicates “Censored.” See “Appendix B – Overlapping Values: Recoded Test Score and Graduation Variables” for institution-specific gradyear values.

Availability

Not available at Southern Methodist.

APPENDIX A – SAT SCORE RECENTERING

COLLEGE BOARD 1995-1996 SAT SCORE RECENTERING

In April 1995, the SAT replaced the former 200-800 verbal and quantitative section test scales with new recentered 200-800 scales. The recentering process created new scales with mean test score much closer to 500. (Prior to the recentering, mean verbal and math section scores were below 500.) Both section and composite recentered scores are higher than equivalent scores that are not recentered.¹

SAT VARIABLES AND RECENTERING FOR THEOP APPLICATION DATA

All institutions report recentered SAT composite scores for 1997 and later years. Seven of the nine THEOP institutions provided SAT data for time periods that extend before and after the recentering. Three of these seven schools (UT Austin, Texas A&M, and UT Pan American) either provided already recentered scores for all years, or provided the verbal and quantitative section scores that are required to generate the equivalent scores on the recentered scale. Recentering scores was done using the College Board concordance table which can be found at:

<http://professionals.collegeboard.com/data-reports-research/sat/equivalence-tables/sat-score>

The remaining four schools (Texas Tech, UT Arlington, UT San Antonio, and Texas A&M Kingsville) provided un-recentered composite scores for the early years and recentered composite scores for later years. The early year scores could not be recentered for these institutions because recentering requires separate verbal and quantitative section scores. SAT scores that could not be recentered are named “sat_not_recenteredR.”

Table 4 summarizes availability and mean values of un-recentered and recentered SAT scores for each THEOP institution.

Table 4. Not-Recentered and Recentered SAT Scores: Availability and Means

	Contains Not Recentered SAT Scores	Years for Not Recentered Scores	Years for Recentered Scores	Mean Not Recentered SAT Scores*	Mean Recentered SAT Scores**
Texas Tech	Yes	1995	1996-2003	958.6	1086.7
UT Arlington	Yes	1994	1995-2002	913.0	1027.2
UT San Antonio	Yes	1990-1996	1997-2004	893.6	968.4
A&M Kingsville	Yes	1992-1995	1996-2002	797.1	914.3
UT Austin	No	N/A	All (1991-2003)	N/A	1185.9
Texas A&M	No	N/A	All (1992-2003)	N/A	1144.7
UT Pan American	No	N/A	All (1995-2002)	N/A	952.0
Rice	N/A	N/A	All (2000-2004)	N/A	1356.3
SMU	N/A	N/A	All (1998-2005)	N/A	1166.7

* across all years not-recentered

** across all years recentered

¹ For more detail on SAT score recentering, see the College Board’s report, “The Recentering of the SAT Scales and Its Effects on Score Distributions and Score Reporting” at:

http://professionals.collegeboard.com/research/pdf/200211_20702.pdf

APPENDIX B – OVERLAPPING VALUES: RECODED TEST SCORE AND GRADUATION VARIABLES

TEST SCORES

Test scores values (satR, actR, testscoreR) are collapsed to maintain applicant confidentiality. Variation in test score distributions across institutions make it impossible to define a single set of non-overlapping categories that eliminate small cell sizes across all institutions. Researchers should be aware that when combining data from two or more institutions, overlapping test score categories may occur for the following variables:

satR, actR, and testscoreR

The applicant file contains an extensive set of test score categories that do not overlap within any individual institution, but that do overlap across institutions.

Table 5 shows score values and categories for the ACT variable (actR).

Table 6 shows score values and categories for SAT (satR) and Testscore (testscoreR) variables.

Table 5. Application Data: ACT Values and Availability across Institutions

<u>value</u>	<u>variable label</u>	UT San Antonio	A&M Kingsville	Texas Tech	UT Arlington	UT Pan American	Southern Methodist	UT Austin	Texas A&M
		1997-2004	1996-2002	1996-2003	1995-2002	1995-2002	1998-2005	1991-2003	1992-2002
1	Less than 16	-	-	-	-	-	-	au	am
2	Less than 17	sa	amk	tt	ar	pa	sm	-	-
3	16-16	-	-	-	-	-	-	au	am
4	17-17	sa	amk	tt	ar	pa	sm	au	am
5	18-18	sa	amk	tt	ar	pa	sm	au	am
6	19-19	sa	amk	tt	ar	pa	sm	au	am
7	20-20	sa	amk	tt	ar	pa	sm	au	am
8	21-21	sa	amk	tt	ar	pa	sm	au	am
9	22-22	sa	amk	tt	ar	pa	sm	au	am
10	23-23	sa	amk	tt	ar	pa	sm	au	am
11	24-24	-	-	tt	ar	pa	sm	au	am
12	24-25	sa	amk	-	-	-	-	-	-
13	25-25	-	-	tt	ar	pa	sm	au	am
14	26-26	-	-	tt	ar	pa	sm	au	am
15	26-36	sa	amk	-	-	-	-	-	-
16	27-27	-	-	-	-	-	sm	au	am
17	27-28	-	-	tt	ar	pa	-	-	-
18	28-28	-	-	-	-	-	sm	au	am
19	29-29	-	-	-	-	-	sm	au	am
20	29-36	-	-	tt	ar	pa	-	-	-
21	30-30	-	-	-	-	-	sm	au	am
22	31-31	-	-	-	-	-	sm	au	am
23	32-32	-	-	-	-	-	sm	au	am
24	33-36	-	-	-	-	-	sm	au	am
.	.	sa	amk	tt	ar	pa	sm	au	am

*Rice does not contain ACT variable

Table 6. Application Data: SAT and Testscore Values and Availability across Institutions

<u>value</u>	<u>variable label</u>	A&M Kingsville 1992- 2002	UT San Antonio 1990- 2004	UT Pan American 1995- 2002	Texas Tech 1996- 2003	UT Arlington 1994- 2002	Southern Methodist 1998- 2005	UT Austin 1991- 2003	Texas A&M 1992- 2002	Rice 2000- 2004
1	Less than 700	amk	sa	pa	-	-	-	-	-	-
2	Less than 750	-	-	-	tt	ar	sm	-	-	-
3	Less than 850	-	-	-	-	-	-	au	am	ri
4	700-720	amk	sa	pa	-	-	-	-	-	-
5	730-750	amk	sa	pa	-	-	-	-	-	-
6	750-790	-	-	-	tt	ar	sm	-	-	-
7	760-770	amk	sa	pa	-	-	-	-	-	-
8	780-790	amk	sa	pa	-	-	-	-	-	-
9	800-820	-	-	-	tt	ar	sm	-	-	-
10	800-830	amk	sa	pa	-	-	-	-	-	-
11	830-850	-	-	-	tt	ar	sm	-	-	-
12	840-860	amk	sa	pa	-	-	-	-	-	-
13	850-890	-	-	-	-	-	-	au	am	ri
14	860-870	-	-	-	tt	ar	sm	-	-	-
15	870-890	amk	sa	pa	-	-	-	-	-	-
16	880-890	-	-	-	tt	ar	sm	-	-	-
18	900-900	-	-	-	tt	ar	sm	-	-	-
19	900-910	amk	sa	pa	-	-	-	-	-	-
20	900-930	-	-	-	-	-	-	au	am	ri
21	910-910	-	-	-	tt	ar	sm	-	-	-
22	920-920	-	-	-	tt	ar	sm	-	-	-
23	920-930	amk	sa	pa	-	-	-	-	-	-
24	930-930	-	-	-	tt	ar	sm	-	-	-
25	940-940	-	-	-	tt	ar	sm	-	-	-
26	940-950	amk	sa	pa	-	-	-	-	-	-
27	940-960	-	-	-	-	-	-	au	am	ri
28	950-950	-	-	-	tt	ar	sm	-	-	-
29	960-960	-	-	-	tt	ar	sm	-	-	-
30	960-970	amk	sa	pa	-	-	-	-	-	-
31	970-970	-	-	-	tt	ar	sm	-	-	-
32	970-990	-	-	-	-	-	-	au	am	ri
33	980-980	-	-	-	tt	ar	sm	-	-	-
34	980-990	amk	sa	pa	-	-	-	-	-	-
35	990-990	-	-	-	tt	ar	sm	-	-	-
36	1000-1000	-	-	-	tt	ar	sm	-	-	-
37	1000-1010	amk	sa	pa	-	-	-	-	-	-
38	1000-1020	-	-	-	-	-	-	au	am	ri
39	1010-1010	-	-	-	tt	ar	sm	-	-	-
40	1020-1020	-	-	-	tt	ar	sm	-	-	-
41	1020-1030	amk	sa	pa	-	-	-	-	-	-
42	1030-1030	-	-	-	tt	ar	sm	-	-	-
43	1030-1040	-	-	-	-	-	-	au	am	ri
44	1040-1040	-	-	-	tt	ar	sm	-	-	-
45	1040-1050	amk	sa	pa	-	-	-	-	-	-
46	1050-1050	-	-	-	tt	ar	sm	-	-	-
47	1050-1060	-	-	-	-	-	-	au	am	ri
48	1060-1060	-	-	-	tt	ar	sm	-	-	-
49	1060-1070	amk	sa	pa	-	-	-	-	-	-
50	1070-1070	-	-	-	tt	ar	sm	au	am	ri
51	1080-1080	-	-	-	tt	ar	sm	au	am	ri
52	1080-1090	amk	sa	pa	-	-	-	-	-	-
53	1090-1090	-	-	-	tt	ar	sm	au	am	ri
54	1100-1100	-	-	-	tt	ar	sm	au	am	ri

<u>value</u>	<u>variable label</u>	A&M Kingsville	UT San Antonio	UT Pan American	Texas Tech	UT Arlington	Southern Methodist	UT Austin	Texas A&M	Rice
55	1100-1140	amk	sa	pa	-	-	-	-	-	-
56	1110-1110	-	-	-	tt	ar	sm	au	am	ri
57	1120-1120	-	-	-	tt	ar	sm	au	am	ri
58	1130-1130	-	-	-	tt	ar	sm	au	am	ri
59	1140-1140	-	-	-	tt	ar	sm	au	am	ri
60	1150-1150	-	-	-	tt	ar	sm	au	am	ri
61	1150-1190	amk	sa	pa	-	-	-	-	-	-
62	1160-1160	-	-	-	tt	ar	sm	au	am	ri
63	1170-1170	-	-	-	tt	ar	sm	au	am	ri
64	1180-1180	-	-	-	tt	ar	sm	au	am	ri
65	1190-1190	-	-	-	tt	ar	sm	au	am	ri
66	1200-1200	-	-	-	tt	ar	sm	au	am	ri
68	1200-1240	amk	sa	pa	-	-	-	-	-	-
69	1210-1210	-	-	-	tt	ar	sm	au	am	ri
70	1220-1220	-	-	-	-	-	-	au	am	ri
71	1220-1230	-	-	-	tt	ar	sm	-	-	-
72	1230-1230	-	-	-	-	-	-	au	am	ri
73	1240-1240	-	-	-	-	-	-	au	am	ri
74	1240-1250	-	-	-	tt	ar	sm	-	-	-
75	1250-1250	-	-	-	-	-	-	au	am	ri
76	1250-1600	amk	sa	pa	-	-	-	-	-	-
77	1260-1260	-	-	-	-	-	-	au	am	ri
78	1260-1270	-	-	-	tt	ar	sm	-	-	-
79	1270-1270	-	-	-	-	-	-	au	am	ri
80	1280-1280	-	-	-	-	-	-	au	am	ri
81	1280-1290	-	-	-	tt	ar	sm	-	-	-
82	1290-1290	-	-	-	-	-	-	au	am	ri
83	1300-1300	-	-	-	-	-	-	au	am	ri
84	1300-1340	-	-	-	tt	ar	sm	-	-	-
85	1310-1310	-	-	-	-	-	-	au	am	ri
86	1320-1320	-	-	-	-	-	-	au	am	ri
88	1330-1330	-	-	-	-	-	-	au	am	ri
89	1340-1340	-	-	-	-	-	-	au	am	ri
90	1350-1350	-	-	-	-	-	-	au	am	ri
91	1350-1390	-	-	-	tt	ar	sm	-	-	-
92	1360-1360	-	-	-	-	-	-	au	am	ri
93	1370-1370	-	-	-	-	-	-	au	am	ri
94	1380-1380	-	-	-	-	-	-	au	am	ri
95	1390-1390	-	-	-	-	-	-	au	am	ri
96	1400-1400	-	-	-	-	-	-	au	am	ri
97	1400-1600	-	-	-	tt	ar	sm	-	-	-
98	1410-1410	-	-	-	-	-	-	au	am	ri
99	1420-1420	-	-	-	-	-	-	au	am	ri
100	1430-1430	-	-	-	-	-	-	au	am	ri
101	1440-1440	-	-	-	-	-	-	au	am	ri
102	1450-1450	-	-	-	-	-	-	au	am	ri
103	1460-1460	-	-	-	-	-	-	au	am	ri
104	1470-1470	-	-	-	-	-	-	au	am	ri
105	1480-1480	-	-	-	-	-	-	au	am	ri
106	1490-1490	-	-	-	-	-	-	au	am	ri
107	1500-1500	-	-	-	-	-	-	au	am	ri
108	1510-1510	-	-	-	-	-	-	au	am	ri
109	1520-1520	-	-	-	-	-	-	au	am	ri
110	1530-1550	-	-	-	-	-	-	au	am	ri
111	1560-1600	-	-	-	-	-	-	au	am	ri
.	.	amk	sa	pa	tt	ar	sm	au	am	ri

GRADUATION YEAR

Variation in distributions as well as variation in years of enrollment data make it impossible to define a single set of non-overlapping categories that eliminate small cell sizes across all institutions. Researchers should be aware that when combining data from two or more institutions, overlapping graduation year categories may occur.

Table 7 shows values and categories for the graduation variable (gradyear).

Table 7. Application Data: Graduation Year Values and Availability across Institutions

value	variable label	UT Austin	Texas A&M	Texas Tech	UT Arlington	UT San Antonio	A&M Kingsville	UT Pan American	Rice
1	1992, 1993, or 1994	-	-	-	-	sa	-	-	-
2	1994 or 1995	-	am	-	-	-	-	-	-
3	1996 or 1997	-	-	-	ar	-	-	-	-
4	1996, 1997, or 1998	-	-	-	-	-	-	pa	-
5	1994	au	-	-	-	-	-	-	-
6	1995	au	-	-	-	sa	-	-	-
7	1996	au	am	-	-	sa	amk	-	-
8	1997	au	am	-	-	sa	amk	-	-
9	1998	au	am	tt	ar	sa	amk	-	-
10	1999	au	am	tt	ar	sa	amk	pa	-
11	2000	au	am	tt	ar	sa	amk	pa	-
12	2001	au	am	tt	ar	sa	amk	pa	-
13	2002	au	am	tt	ar	sa	amk	pa	-
14	2003	au	am	tt	-	sa	amk	pa	ri
15	2004	au	am	tt	-	sa	amk	pa	ri
16	2005	-	am	-	-	-	-	pa	ri
17	2006	-	am	-	-	-	-	-	-
18	2007	-	am	-	-	-	-	-	-
21	Graduated by 2002- actual date not shown	-	-	-	ar	-	-	-	-
22	Graduated by 2004- actual date not shown	au	-	tt	-	sa	amk	-	-
23	Graduated by 2005- actual date not shown	-	-	-	-	-	-	pa	ri
24	Graduated by 2007- actual date not shown	-	am	-	-	-	-	-	-
.c	Censored	au	am	tt	ar	sa	amk	pa	ri
.n	Did not enroll	au	am	tt	ar	sa	amk	pa	ri

APPENDIX C – FIELDS OF MAJOR AND DEPARTMENTS OF MAJOR BY FIELD

Fields in Bold, Majors within field indented below

Agriculture

Architecture

Business

- Accounting
- Business/Management
- Finance
- Marketing/Advertising
- Restaurant and Hotel
- Business, Other

Education

Engineering/Computer Science

- Engineering
- Computer/Information Science

Fine Arts

- Fashion/Interior Design
- Performing Arts
- Textiles
- Visual Arts
- Fine Arts, Other

General Studies

Health

- Allied Health
- Nursing
- Pre-Dentistry
- Pre-Medicine
- Pre-Pharmacy
- Pre-Veterinary Medicine
- Sports/Nutritional Sciences
- Health, Other

Humanities

- English/Literature
- Foreign Language
- Philosophy
- Regional Studies
- Religion
- Humanities, Other

Individualized/Interdisciplinary

Natural/Physical Sciences

- Biology
- Chemistry
- Environmental Science/Studies
- Geography
- Geology
- Mathematics
- Physics
- Zoology
- Natural/Physical Sciences, Other

Other

- Other
- Unknown

Social Sciences

- Anthropology
- Communications
- Criminal Justice
- Economics
- Government/Political Science/Pre-Law
- History
- Human/Child Development
- International Studies
- Journalism
- Policy Studies
- Psychology
- Sociology
- Urban Studies
- Social Sciences, Other

Social Work

Technical/Vocational

Undeclared