Procedures for Developing the Texas Projection Measure Equations

updated March 2011

Background

Texas completed a 2008 pilot study to evaluate two possible approaches to measuring annual student improvement to satisfy requirements under House Bill 1 and Senate Bill 1031. Texas is proposing to use a measure of student growth as a criterion for campuses to meet adequate yearly progress (AYP) for federal reporting purposes in 2009. In addition, Texas will include student growth in the state accountability system for evaluating campuses and districts. The two approaches evaluated in the pilot study included a growth-to-proficiency model and a regression-based projection. The growth-to-proficiency model evaluated in the study was similar to the growth model approved for the United States Department of Education (USDE) growth pilot program for North Carolina. The projection model evaluated in the study was much like the models approved for the USDE growth pilot program in Tennessee and Ohio and was calculated by SAS® and Dr. William Sanders. The two approaches were chosen for the pilot study because they were well matched to the data conditions in Texas, offered the flexibility to potentially satisfy more than one requirement for growth measures, and could be adapted when end-of-course assessments are initiated.

While the pilot study was being conducted, the Texas Select Committee on Public School Accountability convened two groups of district representatives at its April 2008 and August 2008 meetings. These representatives shared information about ways in which their districts had developed and used student growth measures locally. After hearing district testimony at the April meeting, the Select Committee expressed interest in using a regression-based model at the state level, and the Texas Education Agency (TEA) questioned whether the state could implement a model like the one developed by the Dallas Independent School District (DISD), which had been in use since 1992. In summer 2008, TEA evaluated a model similar to the Dallas model, using a few cohorts from the pilot study. A modified projection model similar to that used by Dallas ISD was implemented and ultimately proposed to USDE in spring 2009 for use in calculating AYP for Texas. USDE accepted the Texas proposal in January 2009 and the modified projection model, called the Texas Projection Measure (TPM), was first used in spring 2009 AYP reporting and in the state accountability system in spring 2009.

The purpose of this document is to describe the development and testing of the Texas Projection Measure, a student projection measure much like the Dallas ISD model, using a procedure published by Lissitz, et al. (2006). In particular, this document will describe (1) the procedures used to develop the model formulas to predict grade 8 from grade 7 and to predict grade 11 from grade 10 in mathematics and reading/English language arts using 2006 data predicting 2007 student scores, (2) how the formulas were applied in 2007 to predict 2008 scores, (3) the projection accuracy of the models with the two study cohorts, and (4) how the projection accuracy of these models in 2008 compares with the projection accuracy of the SAS model, which is a more complex regression-based model.

Methods

The procedure used to fit the Texas Projection Measure was one recommended by the Texas Technical Advisory Committee at their July 14–15, 2008, meeting. It consists of two steps and is based on a method described in a paper by Lissitz, et al. (2006). The first step in the process is an ordinary least squares (OLS) multiple regression, which serves to identify variables that statistically relate to measures of achievement. The second step is an analysis of the variability that is due to student clustering within schools, which determines the extent, to which multilevel modeling is justified.

Two sets of analyses were conducted for the development of the grade 8 model formulas. The first set used the 2007 Texas Assessment of Knowledge and Skills (TAKS) grade 8 mathematics (*TAKS_M07*) score as the outcome variable for all analyses. The second set of analyses used TAKS reading (*TAKS_R07*) scores as the outcome variable. From each of these sets of analyses, model formulas were developed. For example, the first set of analyses resulted in the model formulas for projecting to grade 8 mathematics and the second set of analyses resulted in the model formulas for projecting to grade 8 reading. The following section describes the procedures used with the first set of analyses, with the mathematics score as the outcome variable. The procedures were repeated in the second set of analyses for this cohort using the reading score as the outcome variable. Then, the two sets of analyses were conducted for the grade 10 cohort using the mathematics and English language arts scores as the outcomes, respectively.

Procedures

With the 2007 TAKS grade 8 mathematics (*TAKS_M07*) score as the outcome variable, the initial group of student-level predictors entered into the OLS multiple regression included a 2006 TAKS grade 7 reading score (*TAKS_R06*) and a 2006 TAKS grade 7 mathematics score (*TAKS_M06*). These variables were aggregated at the campus level and were included as predictors in the model as well (*MEAN_TAKS_R06*, *MEAN_TAKS_M06*).

The results from the initial OLS regression model indicated that the predictor variables accounted for 66.9% of the variance in the dependent variable, *TAKS_M07*. All predictor variables were significant at an alpha level of 0.05. Thus, all initial predictors were eligible to be included in a multilevel model. The second step of the process involved analyzing the variability of campus-level TAKS scores and the intra-class correlation in order to determine whether a multilevel model was justified. This analysis is conducted using what is known as an unconditional multilevel model. The unconditional model at level 1 can be defined as

$$TAKS\_M07_{ij} = \beta_{0j} + r_{ij} \quad (1)$$

where $TAKS\_M07_{ij}$ represents the 2007 grade 8 mathematics score for individual *i* within school *j*, β_{0j} represents the mean *TAKS_M07* score for school *j*, and r_{ij} represents the residual for individual *i* within school *j*. The variance of $r_{ij} = \sigma^2$.

Level 2 of the unconditional multilevel model can be defined as

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

where γ_{00} is the grand mean of the *TAKS_M07* scores and u_{0j} is the residual for school *j* (i.e., the deviation of school *j* from the grand mean). The variance of u_{0j} = τ_{00}. For the model under consideration, this variance was statistically significant (Z = 26.54, p < .001), meaning there was significant variability in the mean *TAKS_M07* scores among schools. The intra-class correlation is calculated as $\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 5612/(5612+30758) = 0.15$, meaning 15% of the variance of *TAKS_M07* scores is attributable to the effects of students being clustered within schools.

Since the unconditional model indicated variability at the school level, a model with school- and student-level predictors was run. The model may be run twice: once with all the student and school-level predictors indicated by the OLS model, and, if necessary, again with variables omitted that were not statistically significant in the multilevel model. The predictor variables in the model under consideration were all statistically significant, resulting in a final level 1 model

$$TAKS\_M07_{ij} = \beta_{0j} + \beta_{1j}(TAKS\_M06_j) + \beta_{2j}(TAKS\_R06_j) + r_{ij}\ .$$

In multilevel modeling, the level 1 regression coefficients (i.e., the βs) are tested for variability at level 2. If variability at level 2 is indicated, then level 2 predictor variables can be added to the model. For example, variables could be added to the model in an effort to account for the variability in the school means. An example of a level 2 model using mean *TAKS_R06* and mean *TAKS_M06* explanatory variables to account for variability in the school means is presented in Equation 3

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(MEAN\_TAKS\_M06) + \gamma_{02}(MEAN\_TAKS\_R06) + u_{0j} \qquad (3)$$

Equation 3 illustrates the first level 2 model TEA considered; however, complicating factors make such a model difficult to implement in practice at a statewide level. An example of such a complicating factor is cross-classification of students. Cross-classification refers to the fact that students who attend the same school in one year may attend different schools in a subsequent year. A common example of cross-classification occurs when two students from the same elementary school attend different middle schools. Because the model is used to make predictions across schools (e.g., 5^{th} to 8^{th} grade), the school mean explanatory variable was modeled at the student level rather than at the school level. Substituting Equation 2 (rather than Equation 3) into the level 1 equation and adding school mean explanatory variables to level 1 results in the second multilevel prediction equation considered,

$$\begin{aligned} TAKS\_M07_{ij} &= \gamma_{00} + \gamma_{10}(TAKS\_M06) + \gamma_{20}(TAKS\_R06) \\ &+ \gamma_{30}(MEAN\_TAKS\_M06) + \gamma_{40}(MEAN\_TAKS\_R06) + (u_{0j} + r_{ij}) \end{aligned} \qquad (4)$$

This equation, developed using 2006 scores as predictors of 2007 scores, was then used to predict 2008 grade 8 mathematics scores for the 2007 grade 7 cohort. Using equations developed in the prior year would allow equations to be developed and published before they are used in Texas and allowed the state to report projections on students' Confidential Student Reports on the regular reporting schedule. If a student's 2008 predicted score was 2100 (the Met Standard score) or above, that student was classified as meeting growth targets in 2007. If the predicted score was below 2100, that student was classified as not meeting growth targets. Finally, the accuracy of the growth classifications based on predicted scores was assessed by comparing them to the observed

2008 grade 8 results. During this set of analyses, the off-subject school mean predictor coefficient (i.e., *MEAN_TAKS_R06* in Equation 4) was often small in magnitude and was not contributing much to the model in terms of R-square value. Therefore, the prediction equation was refined again to arrive at our final TPM model:

$$\begin{aligned} TAKS\_M07_{ij} &= \gamma_{00} + \gamma_{10}(TAKS\_M06) + \gamma_{20}(TAKS\_R06) \\ &+ \gamma_{30}(MEAN\_TAKS\_M06) + (u_{0j} + r_{ij}) \end{aligned} \qquad (5)$$

The set of classification accuracy analyses described above was then replicated using the final TPM model. The results of the final set of TPM classification accuracy analyses are presented below.

Results

The procedures described above were repeated three times: once to predict grade 8 reading for the grade 7 cohort, once to predict grade 11 mathematics for the grade 10 cohort, and once to predict grade 11 English language arts for the grade 10 cohort. The percentage of variance accounted for by the predictors and the intra-class correlation coefficients are presented in table A1 for the cohorts. The unstandardized regression coefficients and p- values from the multilevel model equations for the two cohorts in both subjects are presented in table A2. The projection accuracy results for all cohorts are contrasted with projection accuracy results from the more complex regression-based EVAAS® projection model and presented in summary form in table A3 and in more detail in tables A4 through A11.

Table A1. *TPM Percent of Variance Accounted For and Intra-Class Correlation Coefficients*

Projection Grade and Subject	Year Formulas Developed	Year Formulas Applied	Year Projection Accuracy Evaluated	Percent of Variance Accounted for by Predictors	Intra-class Correlation Coefficient
Grade 8 Reading	2006	2007	2008	53.6%	0.11
Grade 8 Mathematics	2006	2007	2008	66.9%	0.15
Grade 11 English Language Arts	2006	2007	2008	56.1%	0.16
Grade 11 Mathematics	2006	2007	2008	70.2%	0.18

Table A2. *TPM Unstandardized Regression Coefficients and P-values for Grade 7*

Indicators	Grade 7 Reading		Grade 7 Mathematics	
	Coefficient	p-value	Coefficient	p-value
Constant	187.77	< .0001	57.76	.0007
Student-level variables				
TAKS_R06	0.5927	< .0001	0.1325	< .0001
TAKS_M06	0.2561	< .0001	0.8164	< .0001
School-level variables				
MEAN_TAKS_R06	0.1110	< .0001		
MEAN_TAKS_M06			0.02657	0.0007

Table A2. *TPM Unstandardized Regression Coefficients and P-values for Grade 10*

Indicators	Grade 10 English Language Arts Coefficient	Grade 10 English Language Arts p-value	Grade 10 Mathematics Coefficient	Grade 10 Mathematics p-value
Constant	232.27	< .0001	278.45	< .0001
Student-level variables				
TAKS_R06	0.5729	< .0001	0.1282	< .0001
TAKS_M06	0.2222	< .0001	0.6971	< .0001
School-level variables				
MEAN_TAKS_R06	0.1301	< .0001		
MEAN_TAKS_M06			0.0675	< .0001

Table A3. *Projection Accuracy for the TPM and the EVAAS® Projection Model*

PROJECTION YEAR, GRADE, AND SUBJECT	TEXAS PROJECTION MEASURE			EVAAS® PROJECTION MODEL		
	N	Perfect Agreement Met Standard	Perfect Agreement Did Not Meet Standard	N	Perfect Agreement Met Standard	Perfect Agreement Did Not Meet Standard
2008 Grade 8 Reading	270,670	94	2	269,015	94	2
2008 Grade 8 Mathematics	270,679	73	13	267,540	73	14
2008 Grade 11 English Language Arts	224,547	93	1	225,923	92	3
2008 Grade 11 Mathematics	222,645	79	10	228,110	78	11

Table A4. *Texas Projection Measure* Grade *7 Cohort: Grade 8 Reading*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	5012 (1.85)	3073 (1.14)	8085 (2.99)
Projected Met Growth	8355 (3.09)	254230 (93.93)	262585 (97.01)
Total	13367 (4.94)	257303 (95.06)	270670 (100.00)

Table A5. *EVAAS® Projection Model Grade7 Cohort: Grade 8 Reading*

	Observed Score Did Not Meet Growth	**Observed Score Met Growth**	**Total**
Projected Did Not Meet Growth	5097 (1.89)	3461 (1.29)	8558 (3.18)
Projected Met Growth	7235 (2.69)	253222 (94.13)	260457 (96.82)
Total	12332 (4.58)	256683 (95.42)	269015 (100.00)

Table A6. *Texas Projection Measure Grade 7 Cohort: Grade 8 Mathematics*

	Observed Score Did Not Meet Growth	**Observed Score Met Growth**	**Total**
Projected Did Not Meet Growth	35837 (13.24)	18679 (6.90)	54516 (20.14)
Projected Met Growth	18961 (7.00)	197202 (72.85)	216163 (79.86)
Total	54798 (20.24)	215881 (79.76)	270679 (100.00)

Table A7. *EVAAS® Projection Model Grade 7 Cohort: Grade 8 Mathematics*

	Observed Score Did Not Meet Growth	**Observed Score Met Growth**	**Total**
Projected Did Not Meet Growth	37007 (13.83)	19188 (7.17)	56195 (21.00)
Projected Met Growth	15882 (5.94)	195463 (73.06)	211345 (79.00)
Total	52889 (19.77)	214651 (80.23)	267540 (100.00)

Table A8. *Texas Projection Measure Grade 10 Cohort: Grade 11 English Language Arts*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	3186 (1.42)	766 (0.34)	3952 (1.76)
Projected Met Growth	11536 (5.14)	209059 (93.10)	220595 (98.24)
Total	14722 (6.56)	209825 (93.44)	224547 (100.00)

Table A9. *EVAAS® Projection Model Grade 10 Cohort: Grade 11 English Language Arts*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	6555 (2.87)	3877 (1.70)	10432 (4.57)
Projected Met Growth	8520 (3.74)	209158 (91.69)	217678 (95.43)
Total	15075 (6.61)	213035 (93.39)	228110 (100.00)

Table A10. *Texas Projection Measure Grade 10 Cohort: Grade 11 Mathematics*

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	22127 (9.94)	9046 (4.06)	31173 (14.00)
Projected Met Growth	16437 (7.38)	175035 (78.62)	191472 (86.00)
Total	38564 (17.32)	184081 (82.68)	222645 (100.00)

Table A11. EVAAS® Projection Model Grade 10 Cohort: Grade 11 Mathematics

	Observed Score Did Not Meet Growth	Observed Score Met Growth	Total
Projected Did Not Meet Growth	25484 (11.28)	9550 (4.23)	35034 (15.51)
Projected Met Growth	14976 (6.63)	175913 (77.86)	190889 (84.49)
Total	40460 (17.91)	185463 (82.09)	225923 (100.00)

Discussion

Results from this study indicated that projection accuracy for the Texas Projection Measure was similar to projection accuracy with the more complex EVAAS® projection model for these cohorts. Results from the eight comparisons of accurate prediction percentages in Table A3 illustrated that three were exactly the same, four differed by one percentage point, and one differed by two percentage points. Two of the five comparisons that differed indicated that the Texas model was more accurate than the more complex EVAAS® model. Thus, the Texas model shares some of the advantages of projection models in general and the EVAAS® model in particular. These advantages include that the Texas Projection Measure has evidence supporting its accuracy and reliability, offers the flexibility to be adapted when end-of-course assessments are initiated in the 2011–2012 school year, and would likely increase the accuracy of the calculations of Adequate Yearly Progress for campuses, districts, and the state. Though the Texas model is similar to and produces similar results to the EVAAS® model, it is simpler to implement and uses formulas from the prior year, so that the process for predicting student performance as an indicator of student growth is transparent to the state and can be reported on students' Confidential Student Reports during Texas' regularly-scheduled reporting timeframe. State law requires that schools receive results from the first administration of some TAKS tests within ten working days of receipt of the test materials by the testing contractor.

Because the Texas model uses prior-year equations and is less complex to calculate than the EVAAS® model, it lacks some of the flexibility that the EVAAS® model has in handling missing data. Students must have valid scores in both reading/English language arts and mathematics to be projected in the Texas model. Whereas the Texas model has a slight disadvantage with regard to missing data, it nevertheless has some advantages over the more complex model. For example, the Texas Projection Measure is easy to implement using standard statistical software, so the turnaround time between test completion and projection calculation would be relatively short, and student projection results could be reported on the student reports and used in instructional planning as early as possible. Furthermore, the regression coefficients could be made publicly available so that school and district personnel would be able to calculate projected scores relatively easily. Though the intricacies of the development of the multilevel regression equations may be difficult for stakeholders to understand, the basic idea of using students' current-year test scores to predict future performance is straightforward.

A potential disadvantage of any regression-based model is that the methods underlying the equations are complex and will be difficult for stakeholders who do not have a statistical background to understand. In addition, the accuracy and reliability of regression-based models are likely to decrease the closer a student is to the classification cut score, where small errors can mean the difference between being classified correctly and incorrectly. It has also been shown that projection accuracy decreases as the number of years between testing grade and projection grade increases. See projection accuracy results presented in the January 12, 2009 Texas proposal to the USDE entitled, “Texas Education Agency Growth Model Pilot Application for Adequate Yearly Progress Determinations Under the No Child Left Behind Act,” which is published on the USDE website at the Texas link at the bottom of the page (http://www.ed.gov/admins/lead/account/growthmodel/index.html). Texas analyses to evaluate projection accuracy for all grades and subjects are planned as annual analyses, so that projection accuracy for students at all score points and those being projected one, two, or three years in the future can be documented and monitored over time.

References

Lissitz, R. W., Fan, W., Alban, T., Hislop, B., Strader, D., Wood, C., et al. (March, 2006). *The projection of performance on the Maryland high school graduation exam: Magnitude, modeling and reliability of results*. A paper presented at the National Council on Measurement in Education, San Francisco.