Since students are nested within classrooms, we relied on multilevel modeling techniques to provide more accurate estimates of student and classroom-level effects (see Bryk and Raudenbush, 1992; Bryk, Raudenbush, and Congdon, 1996; Singer, 1998). We estimated a series of models both for all the students in the district and for the survey sample. Specific model specifications for these are described in turn.

ACHIEVEMENT MODELS FOR ALL STUDENTS IN THE DISTRICT

Using district-wide data, the models estimated for both reading and mathematics scores are specified as follows:

**Individual Student Model (Level 1)**

\[ Y_{ij} = \beta_{0j} + \beta_1 \text{(Male)} + \beta_2 \text{(African American)} + \beta_3 \text{(White or Other)} + \beta_4 \text{(Bilingual Education Program)} + \beta_5 \text{(Limited English Proficiency)} + \beta_6 \text{(Free/Reduced Price Lunch Program Eligible)} + \beta_7 \text{(Special Education)} + \beta_8 \text{(Gifted and Talented Education)} + \beta_9 \text{(Age)} + \beta_{10} \text{(Weeks with Class)} + \beta_{11} \text{(1997 TAAS Score)} + \beta_{12} \text{(Missing 1997 TAAS Score)} + r_{ij}, \]
where

- $Y_{ij}$ is the dependent variable, i.e., the TAAS reading or mathematics score for student $i$ in classroom $j$;
- $\beta_{0j}$ is the Level 1 constant term and the average value of the dependent variable in classroom $j$;
- $\beta_{1-12}$ are the Level 1 coefficients for the listed independent variables; and
- $r_{ij}$ is the Level 1 random effect.

**Classroom Context Model (Level 2)**

$$
\beta_{0j} = \gamma_0 + \gamma_1 \text{(Male Teacher)} + \gamma_2 \text{(African American Teacher)} + \gamma_3 \text{(White or Other Teacher)} + \gamma_4 \text{(Teacher's Years of Experience)} + \\
\gamma_5 \text{(Master's Degree or Higher)} + \gamma_6 \text{(Male, $j$)} + \gamma_7 \text{(African American, $j$)} + \gamma_8 \text{(White or Other, $j$)} + \gamma_9 \text{(Bilingual Education Program, $j$)} + \gamma_{10} \text{(Limited English Proficiency, $j$)} + \gamma_{11} \text{(Free/Reduced Price Lunch Program Eligible, $j$)} + \gamma_{12} \text{(Special Education, $j$)} + \gamma_{13} \text{(Gifted and Talented Education, $j$)} + \gamma_{14} \text{(Age, $j$)} + \\
\gamma_{15} \text{(Weeks with Class, $j$)} + \gamma_{16} \text{(1997 TAAS Score, $j$)} + \gamma_{17} \text{(Missing 1997 TAAS Score, $j$)} + \gamma_{18} \text{(Average TEA Rating, 1995 to 1997)} + \\
\gamma_{19} \text{(Years Implementing a NAS Design)} + u,
$$

where

- $\beta_{0j}$ in this model is from the student-level equation above, and it is the average test score in classroom $j$;
- $\gamma_0$ is the constant term;
- $\gamma_{1-19}$ are the Level 2 coefficients for the listed independent variables (any variable with a “$j$” ending is the classroom average of students for that variable); and
- $u$ is a Level 2 random effect.
ACHIEVEMENT MODELS FOR RAND SURVEY SAMPLE

Individual Student Model (Level 1)

At the student level, we use the same model specification as in the district-wide analysis for each of the TAAS reading, mathematics, and Stanford-9 regressions. Within the Stanford-9 model the 1997 TAAS readings scores are used to control for prior student achievement.

\[ Y_{ij} = \beta_0 + \beta_1 \text{(Male)} + \beta_2 \text{(African American)} + \beta_3 \text{(White or Other)} + \beta_4 \text{(Bilingual Education Program)} + \beta_5 \text{(Limited English Proficiency)} + \beta_6 \text{(Free/Reduced Price Lunch Program Eligible)} + \beta_7 \text{(Special Education)} + \beta_8 \text{(Gifted and Talented Education)} + \beta_9 \text{(Age)} + \beta_{10} \text{(Weeks with Class)} + \beta_{11} \text{(1997 TAAS Score)} + \beta_{12} \text{(Missing 1997 TAAS Score)} + r_{ij}, \]

where

- \( Y_{ij} \) is the dependent variable, i.e., the TAAS reading or mathematics score or the Stanford-9 score for student \( i \) in classroom \( j \);
- \( \beta_0 \) is the Level 1 constant term and the average value of the dependent variable in classroom \( j \);
- \( \beta_1-12 \) are the Level 1 coefficients for the listed independent variables;
- \( r_{ij} \) is the Level 1 random effect.

Classroom Context Model (Level 2)

\[ \beta_{0j} = \gamma_0 + \gamma_1 \text{(Male Teacher)} + \gamma_2 \text{(African American Teacher)} + \gamma_3 \text{(White or Other Teacher)} + \gamma_4 \text{(Teacher’s Years of Experience)} + \gamma_5 \text{(Master’s Degree or Higher)} + \gamma_6 \text{(Male.j)} + \gamma_7 \text{(African American.j)} + \gamma_8 \text{(White or Other.j)} + \gamma_9 \text{(Bilingual Education Program.j)} + \gamma_{10} \text{(Limited English Proficiency.j)} + \gamma_{11} \text{(Free/Reduced Price Lunch Program Eligible.j)} + \gamma_{12} \text{(Special Education.j)} + \gamma_{13} \text{(Gifted and Talented Education.j)} + \gamma_{14} \text{(Age.j)} + \gamma_{15} \text{(Weeks with Class.j)} + \gamma_{16} \text{(1997 TAAS Score.j)} + \gamma_{17} \text{(Missing 1997 TAAS Score.j)} + \gamma_{18} \text{(Average TEA Rating, 1995 to 1997)} + \gamma_{19} \text{(Years Implementing a NAS Design)} + \gamma_{20} \text{(Lecture)} + \gamma_{21} \text{(Worksheets)} + \gamma_{22} \text{(Drill on Computational Skills)} + \gamma_{23} \text{(Collaboration)} + \gamma_{24} \text{(Factors Hindering Student
Achievement) + \gamma_{25}(Principal Leadership) + \gamma_{26}(Quality of Professional Development) + \gamma_{27}(Reform) + u,$

where

- $\beta_{0j}$ in this model is from the student-level equation above, and it is the average test score in classroom $j$;
- $\gamma_{0}$ is the intercept term;
- $\gamma_{1-27}$ are the Level 2 coefficients for the listed independent variables (any variable with a “.j” ending is the classroom average of students for that variable); and
- $u$ is a Level 2 random effect.

The multilevel models described in Chapter Six are simple “fixed coefficient” models (Kreft and DeLeeuw, 1998). The coefficients for the Level 1 relationships between student characteristics and achievement scores are held constant across classrooms. There are no cross-level interactions between classroom and student characteristics. Thus, between-classroom differences are limited to differences in intercepts. In other words, the intercept for each classroom is the sum of the overall intercept and the sums of the classroom aggregate variables weighted by the classroom-level regression coefficients, plus random error. The achievement score of each student then is the sum of that student’s classroom intercept and the sum of the student-level coefficients, plus error (Koretz, McCaffrey, and Sullivan, 2001).

Student-level variables were centered on their classroom-level means. Centering is particularly important for these Level 1 variables both for interpretation of the intercept as well as to ensure numerical stability (Bryk and Raudenbush, 1992, p. 25). Without centering, one would interpret the intercept as the expected outcome for a student in classroom $j$ who has a value of zero on all the Level 1 predictor variables, which often does not have meaning. With centering, one can interpret the intercept as the expected outcome for a student in classroom $j$ who is at the mean of all the predictor variables. Dummy variables were also group-mean centered; thus the intercept term is the adjusted mean outcome in classroom $j$, adjusted for differences among units in the percentages of students with various characteristics. The group (classroom)
means of the student-level variables were then included in the Level 2 model to avoid the assumption that the effects of variations in classroom means equal the effects of deviations within classrooms (Bryk and Raudenbush, 1992). This specification also makes the overall model’s coefficients straightforward estimates of the within- and between-classroom effects.¹

¹Cohen et al. (1998) point out the importance of reintroducing the classroom means when the Level 1 variables are group-mean centered. They argue that a model in which the (Level 1) variables are centered on their classroom means sheds an important piece of information: the classroom mean of the variable. Specifically, they write, “When the analyst fails to re-introduce this source of systematic variation appropriately elsewhere in the model, he or she posits that the actual value of the centered does not influence the outcome, only the relative value (that is, relative to the school [or classroom in this context] mean)” (pp. 18-19).