2010-11 Beta Growth Model for Educator Evaluation Technical Report

New York State Education Department

AIR®

AMERICAN INSTITUTES FOR RESEARCH®

Table of Contents

INTRODUCTION 1
DATA 2
Test Scores 2
Teacher-Student-Course Linkages 3
Student and School Characteristics 4
MODEL 5
Covariate Adjustment Model 6
Accounting for Measurement Error in the Predictor Variables 6
Replacing $\boldsymbol{U'\Omega - 1U}$ with Its Expectations 8
Specification for New York Growth Model 9
Student Growth Percentiles 10
Interpolating Standard Errors of Measurement at the Lowest and Highest Obtainable Scale Scores (LOSS and HOSS) 13
Interpolation Procedure for Conditional Standard Errors of LOSS and HOSS 13
Mean Growth Percentiles 14
Aggregation of Growth Percentiles 15
REPORTING 15
Minimum Sample Sizes for Reporting 16
Performance Categories 17
RESULTS 17
Model Fit Statistics 18
Student Growth Percentiles 19
Mean Growth Percentiles 19
Precision of the Mean Growth Percentiles 20
Impact Data Results 23
CONCLUSION 35
REFERENCES 36

List of Appendices

APPENDIX A. TASK FORCE AND TECHNICAL ADVISORY COMMITTEE MEMBERS 37
APPENDIX B. DATA PROCESSING RULES AND RESULTS 38
APPENDIX C. COURSE CODES USED IN ANALYSIS 46
APPENDIX D. ATTRIBUTION RULES 47

APPENDIX E. DEMOGRAPHIC VARIABLE DEFINITIONS 50
APPENDIX F. DESCRIPTIVE RESULTS 50
APPENDIX G. MODEL VARIANCE COMPONENTS 53
APPENDIX H. MODEL COEFFICIENTS 56
APPENDIX J. ADDITIONAL IMPACT DATA (TEACHER LEVEL BY SUBJECT AND GRADE, AND SCHOOL LEVEL) 66

List of Figures

Figure 1: Conditional Standard Error of Measurement Plot, Grade 8 ELA 7
Figure 2: Sample Growth Percentile from Mixed Model 12
Figure 3: Sample Growth Percentile from Mixed Model 12
Figure 4: Distribution of Teacher MGP by Grade and Model 20
Figure 5: Distribution of Principal MGP by Model 20
Figure 6: Overall MGP with 95% Confidence Interval Based on Random Sample of 100 Teachers 21
Figure 7: Overall MGP with 95% Confidence Interval Based on Random Sample of 100 Principals 22
Figure 8: Relationship of Teacher MGP Scores to Percent of ELL Students 25
Figure 9: Relationship of Teacher MGP with Percent SWD in Class 26
Figure 10: Relationship of Teacher MGP Scores to Percent of Economically Disadvantaged Students 27
Figure 11: Relationship of Teacher MGP Scores to Mean Prior ELA Scores 28
Figure 12: Relationship of Teacher MGP Scores to Mean Prior Mathematics Scores 29
Figure 13: Relationship of Principal MGP Scores to Percent of ELL Students 31
Figure 14: Relationship of Principal MGP with Percent SWD in School 32
Figure 15: Relationship of Principal MGP Scores to Percent of Economically Disadvantaged Students 33
Figure 16: Relationship of Principal MGP Scores to Mean Prior ELA Scores 34
Figure 17: Relationship of Principal MGP Scores to Mean Prior Mathematics Scores 34

List of Tables

Table 1: Growth Score Ranges and Categories 17
Table 2: R-Square Values by Grade and Subject 18
Table 3: Correlation between SGP and Prior Year Scale Score 19
Table 4: Percent of Educator MGPs above or below mean at 95% Confidence Level 22

Table 5: Mean Standard Errors, Standard Deviation, and Value of ρ for Adjusted and Unadjusted Models by Grade (teachers only) 23
Table 6: Teacher MGP Correlated with Class Characteristics 24
Table 7: Principal MGP Correlated with Class Characteristics.....30

INTRODUCTION

New York Education Law §3012-c requires a new performance evaluation system for classroom teachers and building principals in New York State. Under the new law, New York State will differentiate teacher and principal effectiveness using four rating categories: Highly Effective, Effective, Developing, and Ineffective (HEDI). Education Law §3012-c(2)(a) requires annual professional performance reviews (APPRs) to result in a single composite teacher or principal effectiveness score, which incorporates multiple measures of effectiveness. The results of the evaluations will be a significant factor in employment decisions, including but not limited to promotion, retention, tenure determinations, termination, and supplemental compensation. It will also affect decisions regarding teacher and principal professional development (including coaching, induction support, and differentiated professional development).

For 2011–12, 20 percent of a teacher or principal's evaluation will be based on student growth on State assessments (for those educators with grades 4-8, ELA and Mathematics), with another 20 percent being based on other locally-selected measures of student achievement, identified through collective bargaining. The remaining 60 points will be based on "other" measures of educator effectiveness, selected by each district through collective bargaining, including classroom observations of teachers.

For the 2011–12 school year, Education Law §3012-c requires the use of student growth data for the first 20% and the statute contemplates a "value-added" model for use in the 2012–13 school year and beyond (increasing the first 20% to 25%). The statute requires that the value-added model be approved by the Board of Regents after consultation with an advisory committee. For the 2011-2012 school year, New York committed to the use of a student growth model that would generate student growth percentile (SGP) scores based on data from the New York State testing program in English Language Arts (ELA) and Mathematics in grades 3 through 8. These SGP scores are aggregated to the teacher and school levels and used for educator evaluation. The student growth model used in 2011-2012 takes into consideration up to three years of prior test history and three student characteristics: poverty, disability status and status as an English Language Learner (ELL). As such, students are compared to similar students i.e., their growth is measured relative to that of other students with similar prior scores and characteristics.

The statute then authorizes districts to use measures derived from the State growth model other than those used for the growth subcomponent for the other 20 percent ("locally selected measures"). For teachers whose students do not take State assessments in grades 4 through 8 in ELA or Mathematics, alternative measures of student learning growth will be developed following State guidance around a student learning objectives process. Results from the growth model will also be incorporated into the State's metrics used for school accountability as part of New York's Elementary and Secondary Education Act (ESEA) waiver.

Because New York's law and regulations envision transitioning to a value-added model, the specific form of the model implemented to develop SGP scores was selected to enable a smooth transition to such a model. Specifically, SGPs are generated from a statistical model know as a mixed random effects model that is described below. Three prior years of test scores are used as

predictor variables, along with a set of measured characteristics for students. The model is an error-in-variables (EiV) regression that accounts for measurement error in the prior test score variables and in the dependent variable.

The models described in this report were selected and developed based on technical and data considerations and on the recommendations of the New York State Education Department (NYSED) as well as members of a statewide advisory committee, (known as the Regents Task Force) including teachers, principals, union representatives, district administrators, and others. A list of Task Force members is provided in Appendix A. In addition, a technical advisory committee reviewed the student growth model for technical accuracy and utility. Technical advisory committee members are also included in Appendix A.

This technical report contains four main sections:

1) **Data.** Description of the data used to implement the student growth model, including data processing rules and relevant issues that arose during processing.
2) **Model.** Statistical description of the model.
3) **Reporting.** Description of reporting metrics and computation of effectiveness scores.
4) **Results.** Overview of key model results aimed at providing information on model quality and characteristics. It is important to note that results presented in this report are based on 2010–11 and prior school years' data. The model will be re-estimated with 2011–12 data when they are available

DATA

To measure student growth and attribute that growth to educators, at least two sources of data are required: student test scores that can be observed over time and information describing how students are linked to schools, teachers and courses (i.e., identifying which teachers teach which students for which tested subjects and which school(s) students attended). In addition, models may also use other information about students and schools, such as student background characteristics, as predictors or controls in the models.

The following section describes the data used for model estimation in New York, including some of the issues and challenges that arose and how they were handled.

Test Scores

New York's student growth model drew on test score data from the statewide testing program in grades 3 through 8 in ELA and Mathematics. These tests are given in the spring. **This report describes how a "beta" model has been constructed using student test results from the 2010–11 school year (and prior years).** The same procedures will be used for the 2011–12 data.

Models were run separately for each grade and subject. Determining which prior achievement scores to use in predicting performance in a particular subject is one key decision to be made in implementing a growth model. The New York State tests at the elementary grade levels include a variety of content aimed at measuring a range of knowledge and skills in Mathematics and ELA. Although the specific content or skills covered may change from year to year, we use test scores

in each subject area as the predictor for that subject area (e.g., Mathematics scores are used to predict Mathematics scores).

In addition to determining which prior test scores to use as predictors, it is also necessary to determine how many years of prior achievement ought to be included as a predictor. In some cases, this may be driven by the availability of data—for example, fourth-grade students typically have only one prior year of scores based on State assessment data. The benefit of including additional years of data is that it may improve the precision of the prediction.

New York's growth models include three prior test scores in the same subject area and missing data indicators for prior test scores except the immediate prior test score. For the 2010–11 analyses, data from 2010–11 were used as outcomes, with prior achievement predictors coming from the three years before (going back to 2007–08). Specific tests used vary by grade and subject, and are as follows:

- Grade 4 ELA and Mathematics models use scores from grade 3 in ELA and Mathematics
- Grade 5 ELA and Mathematics models use scores from grades 3 and 4 in ELA and Mathematics
- Grades 6–8 ELA and Mathematics models use scores from grades 3–7 in ELA and Mathematics.

To implement the EiV approach, the American Institutes for Research (AIR) used conditional standard errors published in the technical reports for the assessments for the outcome and prior-year test scores.

Students who repeated or skipped a grade other than the immediate prior grade were included. For example, a grade 5 student with an immediate prior grade 5 score would be excluded, but a grade 5 student with an immediate prior grade 4 score and a second-year prior score also from grade 4 would be included (with a missing flag being used for the second grade 4 score). This approach was taken in order to include as many students as possible.

Students with no valid immediate prior or outcome scores were dropped from analysis – for example, students without a prior year test score or with a prior year test score for the same grade as the current year test score, so that a growth measure cannot be computed for those students. Students whose scores could not be merged over time were dropped from analysis as well. More detail on exclusion rules and results of applying those rules (along with other specifications) is included in Appendix B.

Teacher-Student-Course Linkages

A critical element of growth analysis is the accurate identification of the courses students are taking in which they learn the content and skills covered on the tests used to measure teacher contributions to their learning. Another critical element is identifying who is teaching those courses.

A first step is to identify which courses are considered "relevant" i.e., courses in which instruction is provided that is aligned to the test being used to measure student growth. New York has developed a common set of course codes across the State, and AIR used the courses

identified as "relevant" by the State for analysis. Appendix C provides a list of the course codes used.

New York also provided data files showing student enrollment in courses and teacher assignment to those courses. However, the State is transitioning to the collection of more detailed linkage data. Notably, there was no way to identify co-teachers or push-in/pull-out support teachers in the current data. Thus, students enrolled in relevant courses were attributed to whatever teacher was scheduled for that course.

Some students were not enrolled in any courses identified as relevant but had test scores in relevant subjects. In these cases, students were attributed to schools (principals) but not individual teachers (through assignment to a "pseudo" teacher variable). Specifically, a little over half of the students in grades 4 and 5 in ELA and Mathematics (53% to 56%) were attributed to schools/principals only; about a third in grade 6 in both subjects (33%) and 8th grade Mathematics (37%); and about a quarter in grades 7 and 8 ELA and grade 7 Mathematics (24% to 26%).

Although course section data were available, a decision was made to provide scores only at the "course-" or "subject-" level, meaning that teachers' scores may reflect multiple classrooms of students in the same content area.

For 2010–11, students were attributed to teachers and principals with whom they were linked at the time of the spring test. We were unable to account for movement across schools and teachers during the year. For 2011–12, NYSED has set a minimum number of days that students must be enrolled in courses with teachers or principals to be attributed to them. For 2012–13 and beyond, districts will provide the State with a more fine-grained "instructional linkage" time, and growth or value-added scores will be weighted based on the amount of instructional time provided by each teacher or in each school. Appendix D shows all current attribution rules.

Student and School Characteristics

In the context of educator evaluation, the results of growth models are used to measure the effects of educators on student learning gains, taking into account a student's prior achievement. It is possible, however, that some factors may impact student learning gains above and beyond educators' contributions, or are confounded or entangled with the educators' contributions. For example, learning gains for a student could be influenced by their disabilities beyond what would be expected based only on their prior achievement. Including a measure of disability status in the model makes an adjustment in what the current year score is expected to be, given disability status, relative to other students.

For 2010–11 and 2011–12, NYSED will use models with and without covariates, so that both "unadjusted" and "adjusted" scores can be provided for teachers and principals. Unadjusted scores will be utilized for school accountability and adjusted scores for educator evaluation.

NYSED's regulations permit three specific control variables at the individual student level for inclusion in the model designed to produce adjusted scores, without additional classroom- or school-level variables. These variables are listed below, and were selected after discussion with New York's Task Force and other stakeholders. Additional variables may be included in the

model in future years, including classroom- or school-level variables that may affect the context in which learning occurs. The student-level predictor variables are:

- **ELL status.** ELL status was defined as being coded as a participant in any of three types of language programs for 2010–11 analysis. Students with any of the three program codes were coded as being ELL (modeled as a Y/N variable). Growth models in subsequent years will explore other options for including ELL status. The programs are shown in Appendix E.
- **Students with disabilities (SWD) status.** A Y/N variable was provided to indicate SWD status.
- **Poverty or economic disadvantage (ED).** Economic disadvantage is coded as economically disadvantaged according to New York State's rules related to family income levels and participation in economic support programs. A description is provided in Appendix E.

MODEL

In this section we describe the implementation of the statistical model to measure student growth in New York. We begin with a description of the statistical model and follow with a description of how SGPs are derived from the estimates produced by the model. In addition, we describe how mean growth percentiles (MGPs) and all variance estimates are produced.

At the core of the New York growth model is the production of an SGP. This is a statistic that characterizes the student's current year score relative to other students with similar prior test score histories. For instance, an SGP equal to 75 denotes that the student's current year score is better than 75 percent of the students in the data with prior test score histories and other measured characteristics that are exactly the same. It does not mean that the student's growth is better than that of 75 percent of all other students in the population.

One common approach to estimating SGPs is to use a quantile regression model (Betebenner, 2008). This approach models the current year score as a function of prior test scores and finds the SGP by comparing the current year score to the predicted values at various quantiles of the conditional distribution.

The methods described here do not rely on the quantile regression method for two reasons. First, the typical implementation of the quantile regression models the observed scores and makes no correction for measurement error in the predictor variables or in the outcome variable. It is known that ignoring the measurement error in the predictor variables yields bias in the model coefficients (e.g., Wei & Carroll, 2009). Further complicating the issue, the measurement error in the outcome variable also adds to the bias in a quantile regression (Hausman, 2001), an issue that does not occur with linear regression.

Second, New York may transition to a value-added model based on methods commonly found in the literature, and there is a desire for congruency between the growth model results and the value-added model results. One way to try and achieve this congruency is by using the statistical

techniques that can produce either the currently used SGPs or another estimate more closely related to Bayes estimates.

The model described in this section is designed to account for measurement error in the predictor variables as well as in the outcome variable to yield unbiased estimates of the model coefficients. Subsequently, these model coefficients are used to form a predicted score, which is ultimately the basis for the SGP. Because the prediction is based on the observed score, it is necessary to account for measurement error in the prediction as well. Hence, the model accounts for measurement error in two steps: first in the model estimation and second in forming the prediction.

Covariate Adjustment Model

The statistical model implemented for the State of New York is typically referred to as a covariate adjustment model (McCaffrey et al, 2003) as the current year observed score is conditioned on prior levels of student achievement as well as other possible covariates.

In its most general form, the model can be represented as:

$$\mathrm{y}_{ti} = \mathbf{X}_i\boldsymbol{\beta} + \sum_{r=1}^{L} \mathrm{y}_{t-r,i}\gamma_{t-r} + \sum_{q=1}^{Q} \mathbf{Z}_{qi}\boldsymbol{\theta}_q + e_i$$

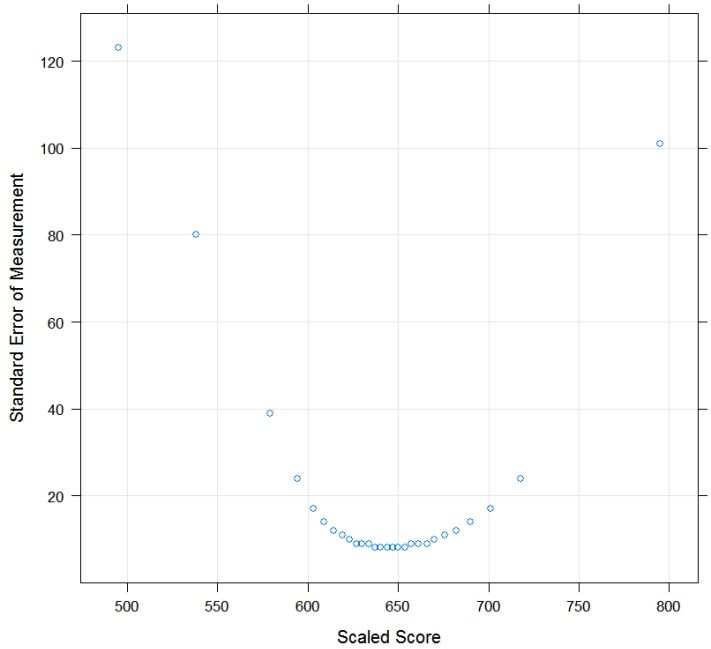
where y_{ti} is the observed score at time *t* for student *i*, $\mathbf{X}_i$ is the model matrix for the student and school level demographic variables, $\boldsymbol{\beta}$ is a vector of coefficients capturing the effect of any demographics included in the model, $y_{t-r,i}$ is the observed lag score at time *t–r* ($r \in \{1,2,\dots,L\}$), $\boldsymbol{\gamma}$ is the coefficient vector capturing the effects of lagged scores, $\mathbf{Z}_{qi}$ is a design matrix with one column for each unit in q ($q \in \{1,2,\dots,Q\}$) and one row for each student record in the database. The entries in the matrix indicate the association between the test represented in the row and the unit (e.g., school, teacher) represented in the column. We often concatenate the sub-matrices such that $\mathbf{Z} = \{\mathbf{Z}_1,\dots,\mathbf{Z}_Q\}$. $\boldsymbol{\theta}_q$ is the vector of effects for the units within a level. In New York, it represents the vector of school or teacher random effects for which we assume $\boldsymbol{\theta}_q \sim N(0, \sigma^2_{\boldsymbol{\theta}_q})$ for each level of q.

Corresponding to $\mathbf{Z} = \{\mathbf{Z}_1,\dots,\mathbf{Z}_Q\}$, we define $\boldsymbol{\theta}' = (\boldsymbol{\theta}'_1,\dots,\boldsymbol{\theta}'_Q)$. In the subsequent sections, we use the notation $\boldsymbol{\delta}' = \{\boldsymbol{\beta}', \boldsymbol{\gamma}'\}$, and $\mathbf{W} = \{\mathbf{X}, \mathbf{y}_{\mathrm{t-1}}, \mathbf{y}_{\mathrm{t-2}}, \dots, \mathbf{y}_{\mathrm{t-L}}\}$ to simplify computation and explanation.

Accounting for Measurement Error in the Predictor Variables

All test scores are measured with error and the magnitude of the error varies over the range of test scores. The standard errors of measurement are referred to as conditional standard errors of measurement (CSEMs) since the variance of a score is heteroscedastic and depends on the score itself. Figure 1 below shows a sample from the grade 8 ELA test in New York.

Figure 1: Conditional Standard Error of Measurement Plot, Grade 8 ELA



Treating the observed scores as if they were the true scores introduces a bias in the regression, and this bias cannot be ignored within the context of a high-stakes accountability system (Greene, 2001). In test theory, the observed score is described as the sum of a true score plus an independent error component, $\boldsymbol{X} = \boldsymbol{X}^* + \boldsymbol{U}$ where $\boldsymbol{U}$ is a matrix of unobserved disturbances with the same dimensions as $\boldsymbol{X}$.

To describe how the model accounts for measurement error, we first re-express the true score regression as:

$$\mathbf{y}_t^* = \mathbf{X}\boldsymbol{\beta} + \sum_{r=1}^{L} \mathbf{y}_{t-r}^* \boldsymbol{\gamma}_{t-r} + \sum_{q=1}^{Q} \mathbf{Z}_q \boldsymbol{\theta}_q + \mathbf{e}$$

We use * to denote the variables without measurement error. For convenience, define the matrices $\mathbf{W} = \{\mathbf{X}, \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{t-L}\}$, $\mathbf{W}^* = \{\mathbf{X}, \mathbf{y}_{t-1}^*, \mathbf{y}_{t-2}^*, \dots, \mathbf{y}_{t-L}^*\}$, and $\boldsymbol{\delta}' = \{\boldsymbol{\beta}', \boldsymbol{\gamma}'\}$. Label the matrix of measurement error disturbances $\boldsymbol{U}$ for disturbances associated with $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{t-L}$, and label the vector of measurement disturbances with the dependent variable, $\mathbf{y}_t$, $\mathbf{v}$, hence $\mathbf{y}_t = \mathbf{y}_t^* + \mathbf{v}$. Let $\boldsymbol{U}$ have the same dimension as $\mathbf{W}$, but only the final L columns of $\boldsymbol{U}$ are non-zero, so $\mathbf{W} = \mathbf{W}^* + \mathbf{U}$. If those disturbances were observed, the parameters $\{\boldsymbol{\delta}', \boldsymbol{\theta}'\}$ can be estimated using Henderson's methods (1950) by solving the following mixed model equations:

$$\begin{pmatrix} \boldsymbol{W}^{*\prime}\boldsymbol{\Omega}^{-1}\mathbf{W}^* & \boldsymbol{W}^{*\prime}\boldsymbol{\Omega}^{-1}\mathbf{Z} \\ \boldsymbol{Z}'\boldsymbol{\Omega}^{-1}\mathbf{W}^* & \boldsymbol{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z} + \mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}'\boldsymbol{\Omega}^{-1}\mathbf{y}^* \\ \boldsymbol{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}^* \end{pmatrix}$$

The matrix $\mathbf{D}$ is made up of Q diagonal blocks, one for each level in the hierarchy. Each diagonal is constructed as $\sigma_q^2 \mathbf{I}_q$ where $\mathbf{I}_q$ is an identity matrix with dimension equal to the number of units

at level q, and σ_q^2 is the estimated variance of the random effects among units at level q. When concatenated diagonally the square matrix $\mathbf{D}$ has dimension $m = \sum_{q=1}^{Q} J_q$.

Two complications intervene. First, we cannot observe $\mathbf{U}$, and second, the unobservable nature of this term, along with the heterogeneous measurement error in the dependent variable, renders this estimator inefficient.

Addressing the first issue, upon expansion we see that:

$$\mathbf{W}^{*\prime}\boldsymbol{\Omega}^{-1}\mathbf{W}^{*} = (\mathbf{W}' - \mathbf{U}')\boldsymbol{\Omega}^{-1}(\mathbf{W} - \mathbf{U}) = \mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{W} - \mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{W} - \mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{U} + \mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U}$$

Since $\mathbf{W} = \mathbf{W}^{*} + \mathbf{U}$, we have $\mathrm{E}(\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{U}) = \mathrm{E}(\mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U})$, $\mathrm{E}(\mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{W}) = \mathrm{E}(\mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U})$, hence $\mathbf{W}^{*\prime}\boldsymbol{\Omega}^{-1}\mathbf{W}^{*} = \mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{W} - \mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U}$. Furthermore, we have $\mathbf{W}^{*\prime}\boldsymbol{\Omega}^{-1}\mathbf{Z} = \mathrm{E}(\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{Z})$, $\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{W}^{*} = \mathrm{E}(\mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{W})$, and $\begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{y}^{*}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}^{*}\end{pmatrix} = \mathrm{E}\begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{y}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}\end{pmatrix}$.

Addressing the second issue, both the right-side and left-side variables in the model equation measured with error contribute to the heteroscedasticity. While the correction $\mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U}$ eliminates the bias due to measurement error, we still do not have an error-free measure of $\mathbf{y}$ for any time period. Therefore, the residual is made up of:

$$\bar{\bar{\mathbf{y}}} - \mathbf{W}'\boldsymbol{\delta} = -\mathbf{U}'\boldsymbol{\delta} + \mathbf{v} + \mathbf{e}\,,$$

where $\bar{\bar{\mathbf{y}}} = \mathbf{y} - \mathbf{Z}\widetilde{\boldsymbol{\theta}}$, $\widetilde{\boldsymbol{\theta}}$ is the conditional mean of the random effects. The residual variance of any given observation is

$$\sigma_{ti}^2 = \sigma_e^2 + \sigma_{v(ti)}^2 + \sum_{r=1}^{L} \delta_{t-r}^2 \sigma_{u,t-r(i)}^2,$$

where $\sigma_{v(ti)}^2$ is the known measurement error variance of the dependent variable for student i at time t. Similarly, $\sigma_{u,t-r(i)}^2$ are the known measurement error variances of r prior test scores. Now, let $\boldsymbol{\Omega}$ be a diagonal matrix of dimension N with diagonal elements σ_{ti}^2.

With the above, we can define the mixed model equations as:

$$\begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{W} - \mathbf{U}'\boldsymbol{\Omega}^{-1}\mathbf{U} & \mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{Z}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{W} & \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z} + \mathbf{D}^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{\delta}\\ \boldsymbol{\theta}\end{pmatrix} = \begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{y}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}\end{pmatrix}$$

Replacing $U'\Omega^{-1}U$ with Its Expectations

As indicated, $\mathbf{U}$ is unobserved and so solving the mixed model equation cannot be computed unless $\mathbf{U}$ is replaced with some observed values. First, the mixed model equations are redefined as:

$$\begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{W} - \mathbf{S} & \mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{Z}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{W} & \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{Z} + \mathbf{D}^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{\delta}\\ \boldsymbol{\theta}\end{pmatrix} = \begin{pmatrix}\mathbf{W}'\boldsymbol{\Omega}^{-1}\mathbf{y}\\ \mathbf{Z}'\boldsymbol{\Omega}^{-1}\mathbf{y}\end{pmatrix}$$

where $\mathbf{S}$ is a diagonal "correction" matrix with dimensions p x p accounting for measurement error in the predictor variables, $p = p_{\mathbf{X}} + L$, and $p_{\mathbf{X}}$ is the column dimension of $\mathbf{X}$.

The matrix **S** is used in lieu of $\mathbf{U}'\mathbf{\Omega}^{-1}\mathbf{U}$ based on the following justification. Recall that we previously defined $\mathbf{\Omega}$ as $\text{diag}(\sigma_{t1}^2, \sigma_{t2}^2, \dots, \sigma_{tN}^2)$ and the matrix of unobserved disturbances is:

$$\boldsymbol{U} = \begin{bmatrix} \mathbf{0}_{p_\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{U}_L \end{bmatrix}$$

where $\mathbf{0}_{p_\mathbf{X}}$ is a matrix of dimension of $p_\mathbf{X}$ with elements of 0, and:

$$\boldsymbol{U}_L = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1L} \\ u_{21} & u_{22} & \dots & u_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \dots & u_{NL} \end{bmatrix}$$

The theoretical result of the matrix operation yields the following symmetric matrix:

$$\mathbf{U}_\mathrm{L}{}'\mathbf{\Omega}^{-1}\mathbf{U}_\mathrm{L} = \begin{bmatrix} \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{i1}^2 & & \dots & \\ \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{i1}u_{i2} & \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{i2}^2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{i1}u_{iL} & \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{i2}u_{iL} & \dots & \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} u_{iL}^2 \end{bmatrix}$$

The theoretical result is limited only because we do not observe u_{ip} since it is latent. However, $E(u_{ip}u_{ip}) = \sigma_{ip}^2$ where σ_{ip}^2 is taken as the conditional standard error of measurement for student i. The theoretical result also simplifies because errors of measurement on different variables are by expectation uncorrelated, $E(u_{ip}u_{ip'}) = 0$ when $p \neq p'$.

Because the conditional standard error of measurement varies for each student i and the off-diagonals can be ignored, let **S** be:

$$\mathbf{S} = \text{diag}\left(0, \dots, 0, \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} \sigma_{u,t-1(i)}^2, \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} \sigma_{u,t-2(i)}^2, \dots, \sum_{i=1}^{N} \frac{1}{\sigma_{ti}^2} \sigma_{u,t-L(i)}^2\right)$$

where $\sigma_{u,j(i)}^2$ denotes the measurement error variance for the jth, $j = (1, 2, \dots L)$, variable measured with error.

Specification for New York Growth Model

The preceding section provides details on the general modeling approach and specifically how measurement error is accounted for in the model. The exact specification for the New York model is described as:

$$y_{gi} = \mu + \sum_{l=1}^{K} \beta_l y_{g-r,i} + \sum_{s=1}^{M} \tau_s m_{si} + \sum_{q=1}^{J} \gamma_q x_{qi} + \theta_m + \theta_{m(j)} + \varepsilon_i$$

where y_{gi} is the current year test scaled score for student i in grade g; μ is the intercept; β_l is the set of coefficients associated with the three prior test scores; τ_s is the set of coefficients associated with the missing variable indicators; γ_q is the set of coefficients associated with the student-level measured characteristics (which include binary variables for poverty, SWD, and ELL status); and θ_m, $\theta_{m(j)}$, and ε_i are the school, teacher, and student random effects.

There are only two missing variable indicators for the prior scores. This occurs because students are required to have the immediate prior score, $y_{g-1,i}$, in order to be included in the model. However, the values for $y_{g-2,i}$ or $y_{g-3,i}$ can be missing. If so, they are accounted for via the missing flag as:

$$\tau = \begin{cases} 1 \text{ if missing} \\ 0 \text{ otherwise} \end{cases}$$

As noted earlier, the model is implemented separately for each grade and subject. When the model includes the covariates for poverty, SWD, and ELL status, it is referred to as the "adjusted" model. The "unadjusted" model is simply a special case of this model which it does not contain those fixed effects, but everything else is the same.

Student Growth Percentiles

The model described in the section above yields unbiased estimates of the fixed effects. For purposes of the growth model, a predicted value and its variance for each student are required to compute the SGPs as:

$$\text{SGP}_\text{i} = \Phi\left(\frac{y_i - \hat{y}_i}{\sqrt{\sigma^2_{yf,i}}}\right)$$

where y_i is the observed value of the outcome variable and $\hat{y}_i = \boldsymbol{w}'\hat{\boldsymbol{\delta}}$ where $\boldsymbol{w}'$ is the ith row of the model matrix $\boldsymbol{W}$ and the notation σ^2_{yfi} is used to mean the variance of a the predicted value of y for the ith student.

Here the regression is of the form

$$\boldsymbol{y} = \boldsymbol{W\delta} + \epsilon$$

Where

$$\epsilon \sim N(0, \sigma_e^2)$$

The classic variance of a predictor is, for this case,

$$\sigma^2_{yf,i} = [1 + \boldsymbol{w_i}'(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{w_i}]\hat{\sigma}_e^2$$

Where $\hat{\sigma}_e^2$ is the standard error of the regression. However, in this case, we wish to make two refinements. The first is to use the actual variance on y_i, called σ_{yi}^2 , rather than the population variance on y_i, called $\bar{\sigma}_{yi}^2$, which is already included in $\hat{\sigma}_e^2$. This is done by subtracting the population variance and adding back the individual variance. Thus the variance on the predictor becomes

$$\sigma_{yf,i}^2 = [1 + \boldsymbol{w_i}'(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{w_i}]\left[\sigma_e^2 + \left(\sigma_{yi}^2 - \bar{\sigma}_{yi}^2\right)\right]$$

The second refinement is to replace the population variance in $\boldsymbol{w_i}$, called $\overline{\boldsymbol{\Sigma}}$, with the individual variance in $\boldsymbol{w_i}$, called $\boldsymbol{\Sigma}_\mathbf{i}$. This is done in the same way as the variance in y_i, so the variance estimate is now

$$\sigma_{yf,i}^2 = [1 + \boldsymbol{w_i}'(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{w_i}]\left[\sigma_e^2 + \left(\sigma_{yi}^2 - \bar{\sigma}_{yi}^2\right) + (\boldsymbol{\beta}'\boldsymbol{\Sigma}_\mathrm{i}\boldsymbol{\beta} - \boldsymbol{\beta}'\overline{\boldsymbol{\Sigma}}\boldsymbol{\beta})\right]$$

There is then a predicted value for each student used to compute the SGP. However, that prediction is based on the estimates of the fixed effects that were corrected for measurement error but based on the error-prone observed score in the vector $\boldsymbol{w}$.

Figure 2 below provides an illustration of how the SGPs are found from the previously-described approach. The illustration considers only a single predictor variable although the concept can be generalized to multiple predictor variables as presented above.

For each student, we find a predicted value conditional on their observed prior scores and the model coefficients. To illustrate the concept, assume we find the prediction and its variance but do not account for the measurement error in the observed scores used to form that prediction. We would form a conditional distribution around the predicted value and find the portion of the normal distribution that falls below their observed score. This is equivalent to:

$$\mathrm{SGP_i} = \int_{-\infty}^{y_i} f(x)dx$$

with $f(x) \sim N(\hat{y}_i, \sigma_{yfi}^2)$ although this is readily accomplished via the cumulative normal distribution function, $\Phi(.)$.

Figure 2: Sample Growth Percentile from Mixed Model

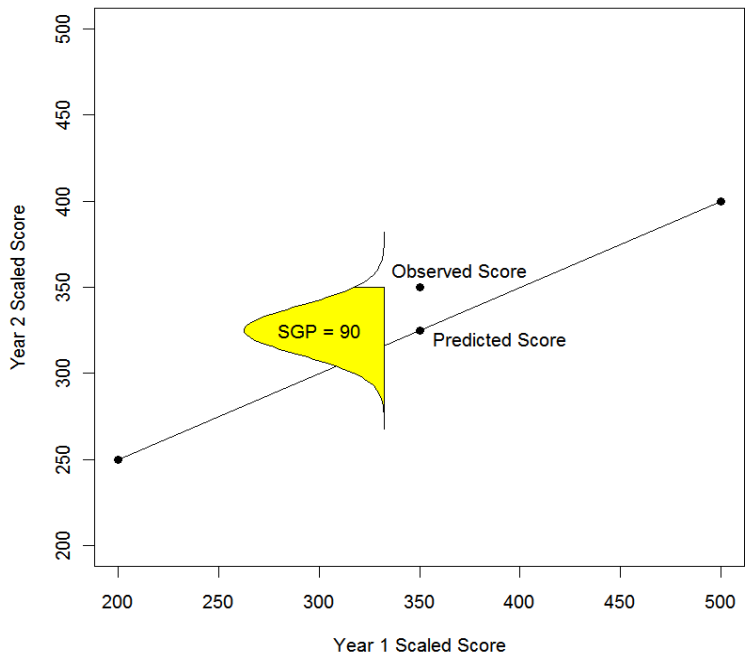
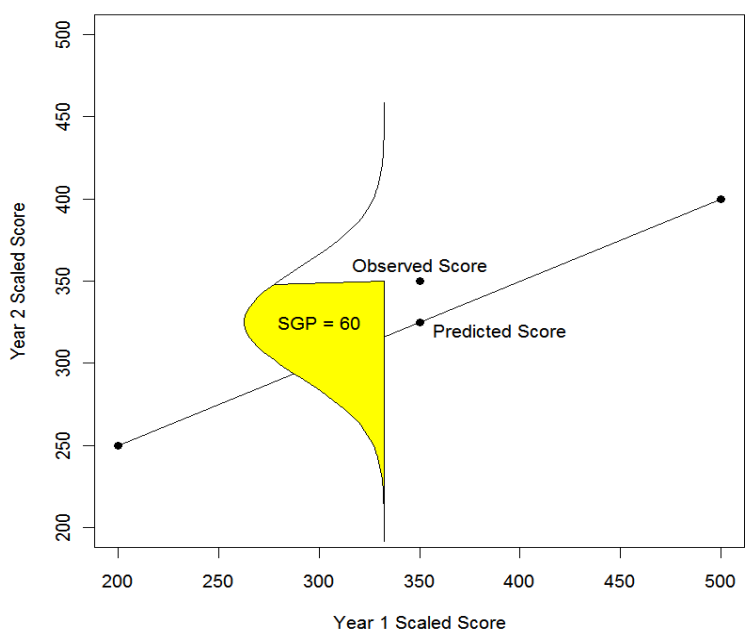


Figure 3 below illustrates the same hypothetical student as above. Note that the observed score and predicted value are exactly the same. However, the prediction variance is larger than in Figure 2 above. As a result, when we integrate over the normal from $-\infty$ to y_i the SGP is 60 and not 90 as in the example above. This occurs because the conditional density curve has become more "spread out," reflecting less precision in the prediction.

Figure 3: Sample Growth Percentile from Mixed Model



Interpolating Standard Errors of Measurement at the Lowest and Highest Obtainable Scale Scores (LOSS and HOSS)

The linear model used to produce student-level predictions $\hat{\boldsymbol{y}}_{\boldsymbol{i}}$ can cause these predictions to fall outside the boundaries of the defined score scale. Let the floor or ceiling in the data be denoted as $\boldsymbol{\eta_f}$ and $\boldsymbol{\eta_c}$, respectively. It is therefore possible that $\hat{\boldsymbol{y}}_{\boldsymbol{i}} < \boldsymbol{\eta_f}$ or $\boldsymbol{\eta_c} < \hat{\boldsymbol{y}}_{\boldsymbol{i}}$. However, the observed score can never fall outside these bounds.

When a prediction falls outside the boundaries of the score scale, it can cause bias in the statistics used to characterize a student, teacher, or school. This phenomenon seems to occur as a result of the large conditional standard errors of measurement at the extreme scores, $\boldsymbol{csem}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}})$. The procedure below is implemented to deal with these large standard errors.

Interpolation Procedure for Conditional Standard Errors of LOSS and HOSS

Interpolate new conditional standard errors of measurement from the following kth degree polynomial regression:

$$\boldsymbol{y_i} = \boldsymbol{\mu} + \sum_{\boldsymbol{k=1}}^{\boldsymbol{M}} \boldsymbol{\beta_k \theta_i^k} + \boldsymbol{e_i}$$

where $\boldsymbol{y_i}$ is $\boldsymbol{csem}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}})^{\boldsymbol{2}}$ and $\boldsymbol{\theta}$ is the observed score for the ith student. The square root of the fitted values will then be used in lieu of the CSEM:

$$\hat{\boldsymbol{y}}_{\boldsymbol{i}} = \hat{\boldsymbol{\mu}} + \sum_{\boldsymbol{k=1}}^{\boldsymbol{M}} \hat{\boldsymbol{\beta}}_{\boldsymbol{k}} \boldsymbol{\theta_i^k}$$

Implementation

Implement the linear regression and subsequently the growth model using the following steps:

1. Subset the data and include all scores except the values at $\boldsymbol{\eta_f}$ and $\boldsymbol{\eta_c}$.
2. Set M = 2 and run linear regression.
3. Using the coefficients from the regression, interpolate new values of $\boldsymbol{csem}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}})^{\boldsymbol{2}}$ for the scores at the LOSS and HOSS.
4. Use $\boldsymbol{csem}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}})$ from step 3 in lieu of the standard error of the LOSS or HOSS in the test score data.
5. Run the growth model.
6. Verify that $\boldsymbol{\eta_f} < \hat{\boldsymbol{y}}_{\boldsymbol{i}} < \boldsymbol{\eta_c}$ for all i.
7. If the inequality in step 6 is not true, return to step 2 and increase to M = M + 1.
8. Repeat until the inequality in step 6 is true.

Mean Growth Percentiles

Once the SGPs are estimated for each student, group-level (e.g., teacher-level) statistics can be formed that characterize the typical performance of students within a group. The NYSED growth model technical advisory committee recommended using the mean SGP when providing educator scores. Hence, group-level statistics are expressed as the mean SGP within a group. This is referred to as the MGP.

For each aggregate unit j ($j \in \{1,2,\dots,J\}$), such as a class, the interest is a summary measure of growth for students within this group. Within group j we have $\{SGP_{j(1)}, SGP_{j(2)}, \dots, SGP_{j(N)}\}$. That is, there is an observed SGP for each student within group j.

Then the MGP for unit j is produced as:

$$\theta_j = \text{mean}(SGP_{j(i)})$$

Like all statistics, the MGP is an estimate and it has a variance term. For New York, AIR provides the following measures of variance for the MGP.

The analytic standard error of the MGP is computed within unit j as:

$$\text{se}(\theta_j) \approx \frac{\text{sd}(\text{SGP}_{ij})}{\sqrt{N_j}}$$

Where $\text{sd}(\text{SGP}_{ij})$ is the sample standard deviation of the SGPs in group j and N is the number of students in group j.

The analytic standard error has two particular limitations. First, MGPs are bounded between 1 and 99; hence the standard errors cannot be used to form confidence limits around the MGP because the confidence limits must be asymmetric. Second, the standard errors do not account for potential non-normality of the distribution of the SGPs. To account for these issues, a bootstrap procedure was implemented.

For each student, we compute an SGP based on

$$\text{SGP} = \Phi\left(\frac{y_i - \hat{y}_i}{\sqrt{\sigma^2_{xf,i}}}\right)$$

The following bootstrap procedure is used:

1. Produce SGP= $\Phi\left(\frac{y_i - \hat{y}_i}{\sqrt{\sigma^2_{yf,i}}}\right)$
2. For each teacher j, sample n_j students with replacement where n_j is the total number of students linked to teacher j.
3. Compute the MGP for teacher j as

a. $\mathrm{MGP_j} = \frac{1}{n_j}\sum_{i=1}^{n_j} \mathrm{SGP_i}$

b. Store the value computed in step 4(a).

4. Repeat steps 2 through 4 five hundred times.
5. Find the bootstrap confidence intervals and standard errors from the vector $\{\mathrm{MGP}_j^1, \mathrm{MGP}_j^2, \ldots, \mathrm{MGP}_j^{500}\}$ where the superscript denotes the qth iteration of the bootstrap. The confidence interval is a range from the 2.5th percentile to the 97.5th percentile of the replicated MGP values.

Aggregation of Growth Percentiles

Teachers in some cases and principals will have students from different grades and with results from different tested subjects. For evaluation purposes, there is a need to aggregate these SGPs and form summary measures.

Because the SGPs are expressed as percentiles, they are free from scale-specific inferences and can be easily combined since percentiles are all on the same scale. For any aggregate-level statistics to be provided (in this case, MGPs), then, we simply pool all SGPs of relevant students and find the mean of the pooled SGPs. Variances of these SGPs are found using the same methods described above. More detail on what scores will be provided follows.

REPORTING

The main metrics generated based on the student growth model include:

- **Unadjusted MGPs.** Unadjusted MGPs are created by simply finding the mean of the SGPs at the relevant level produced by a growth model which does *not* include any contextual variables.
- **Adjusted MGPs (and upper and lower limits based on confidence intervals for these adjusted MGPs).** Adjusted MGPs are created by simply finding the mean of the SGPs at the relevant level produced by a growth model which *does* include three contextual variables at the individual student level (ELL status, SWD status, and economic disadvantage).
- **Percent of students above the median**. Percent of students above the median is found as the percent of students above the State median in the subject and grade using the nominal student SGP from the adjusted model. Note that a teacher-specific growth score computed by NYSED based on the percentage of students earning a State-determined level of growth was required by 2012 New York State legislation to be constructed for use as a locally-selected measure if districts negotiate it locally, as a "different measure" of student growth based on State assessments. It will be provided to all districts, but will be of consequence for evaluation purposes only where districts have locally agreed to it in collective bargaining.

- **Number of student scores included in estimates at each reported level.** This element gives the count of student scores used to compute the adjusted or unadjusted MGP at any given level.
- **Growth rating (HEDI classification for growth subcomponent).** Based on an overall MGP for a teacher or principal across grades and subjects, the growth rating will describe the educator's performance category (Highly Effective, Effective, Developing, or Ineffective). Information on how these classifications are derived is available from NYSED.
- **Growth score.** Based on an overall MGP for a teacher, a growth score of 0–20 will give a finer-grained score within each State growth performance category. Note that this range may change to 0-25 in subsequent years should a value-added model be implemented.

Currently, no data about individual students are planned to be provided as part of this initiative although student-level growth data will be provided to districts as part of NYSED's commitments under the ESEA waiver and its institutional accountability reporting.

MGP results for individual educators will be broken down by subject and grade in addition to overall. Overall MGPs (that is, MGPs based on all students who meet the criteria for a given teacher or school) will also be broken down by several different subgroups (ELL, SWD, economically disadvantaged, low- and high-achieving). ELL, SWD, and economically disadvantaged subgroups will be created using the same rules used for the model (i.e., yes/no for all variables). Subgroups of low- and high-achieving students will be based on looking at performance levels. Specifically, a student who is at level 1 in either subject will be considered low-achieving; a student who is at level 4 in either subject will be considered high-achieving; students at levels 2 in one subject and 3 in the other are not included. Growth subcomponent ratings (HEDI categories and associated evaluation points) are not broken down by any of these sub-categories.

Minimum Sample Sizes for Reporting

One question to be addressed when providing teacher or principal results is whether some of them should be suppressed because the number of students used to create the MGP would jeopardize either personal student information or the quality or stability of the estimate. Setting no (or a low) minimum sample size will result in the greatest number of teachers and principals receiving information; on the other hand, the quality of the information they receive may be poor.

When a larger number of student tests are aggregated together to estimate an MGP, the estimate is more accurate in terms of its estimated standard error. Thus smaller classes have less accurate estimates. To maintain the quality of the estimates, it is therefore desirable to specify a minimum number of students below which the MGP is not reported. For 2010–11, a threshold of 16 student scores was chosen to minimize the number of teachers dropped from the analyses while maintaining the statistical accuracy of estimates.

Performance Categories

Teacher and principal performance (effectiveness) scores are based on their **overall** MGP scores (that is, the results for all the students for whom a teacher or principal is responsible in Mathematics and/or ELA across grades). As noted, to create these overall MGP scores, we simply find the mean of SGPs for all students who are linked to a particular teacher or school.

For 2011–12, teachers and principals are to be assigned to four performance categories, and earn points within each category, as shown in Table 1.

Table 1: Growth Score Ranges and Categories

HEDI Classification	Definition	HEDI Score Range
Ineffective	Results are well-below State average for similar students (or district goals if no State test).	0–2
Developing	Results are below State average for similar students (or district goals if no State test).	3–8
Effective	Results meet State average for similar students (or district goals if no State test).	9–17
Highly Effective	Results are well-above State average for similar students (or district goals if no State test).	18–20

A description of the method used to assign teachers and principals to these categories is available from NYSED.

RESULTS

This section provides an overview of the results of model estimation using 2010-11 as the most recent test year available and the available information about teacher/student data linkages. A pseudo R-squared statistic and summary statistics characterizing the SGPs, MGPs, and their precision provide an overview of model fit. Note that this section focuses on teacher-level results, although additional information on principal/school-level results is available in the appendices.

Appendices G and H provide model parameters, including model coefficients and variance components by grade and subject.

Model Fit Statistics

The R-square is a statistic commonly used to describe the goodness-of-fit for a regression model. Because the model implemented here is a mixed model and not a least squares regression, we refer to this as a pseudo R-square. Table 2 presents the R-square values for each grade and subject, computed as the squared correlation between the fitted values and the outcome variable.

Table 2: R-Square Values by Grade and Subject

Subject	Grade	Unadjusted Model	Adjusted Model
ELA	4	0.315	0.369
	5	0.568	0.573
	6	0.568	0.581
	7	0.577	0.586
	8	0.560	0.579
Math	4	0.416	0.472
	5	0.625	0.631
	6	0.630	0.636
	7	0.650	0.651
	8	0.607	0.610

The R-squared values in grade 4 for both subjects are lower than they are in the other grades. This may be due to the fact that in this grade the model conditions on only one prior test score, while other grades use two or three prior test scores. Notably, however, the R-squared value in grade 5 is comparable in magnitude to the R-squared values in grades 6 through 8, and that model only conditions on two prior scores. These models with lower R-squared values in grade 4 will produce estimated MGPs with higher noise. This will be reflected in the estimated standard errors and corresponding confidence intervals which are used to classify educators.

Student Growth Percentiles

The SGPs describe a student's current year score relative to other students in the data with similar prior academic histories and measured characteristics. Table 3 provides the correlation between the prior-year scaled score and the SGP for each grade and subject.

Table 3: Correlation between SGP and Prior Year Scale Score

Grade	ELA	Math
4	–0.17	–0.02
5	–0.18	–0.07
6	–0.07	–0.12
7	–0.02	–0.15
8	–0.08	0.03

Mean Growth Percentiles

As described earlier in this report, teachers' MGPs are the aggregate educator-level statistics, computed as the mean SGP for all students associated with a roster (teacher), or school (principal). In this section we provide descriptive statistics on the MPG, the average of SGP within the roster / school.

At the teacher level, about 60 percent of teachers have results for students in only one subject (either Mathematics or ELA), and in those instances their MGP is subject-specific. For teachers with results for students in both subjects, the MGP is combined across subjects.

Figure 4 below provides a histogram of the combined teacher MGPs for the adjusted model (includes ELL, SWD, and ED) and the unadjusted model (does not include ELL, SWD, and ED). In both instances, the results are normally distributed. Additional results are shown in Appendix F.

Figure 4: Distribution of Teacher MGP by Grade and Model

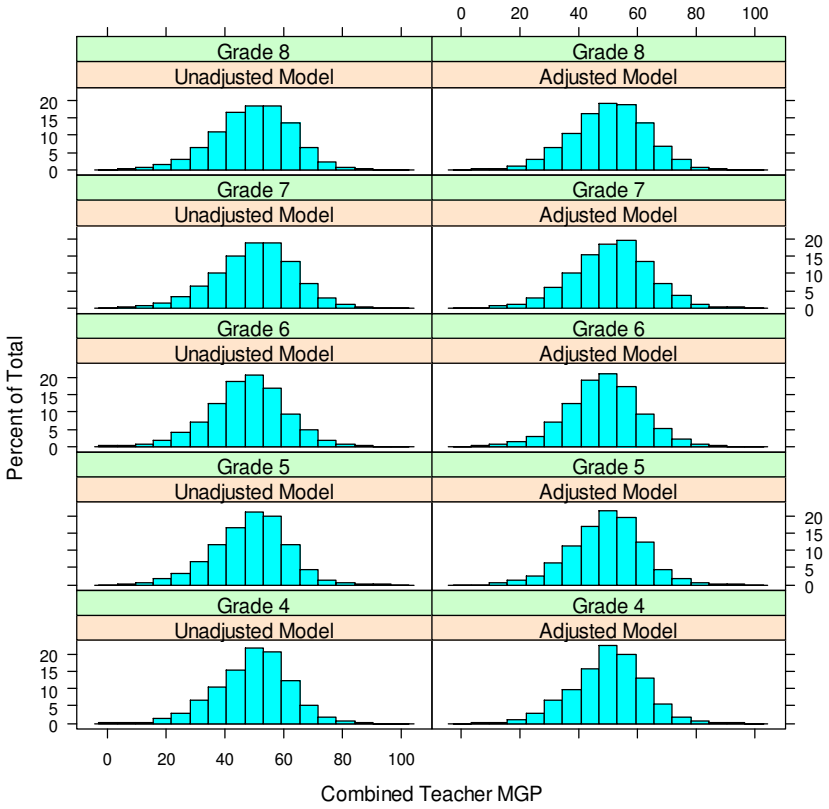
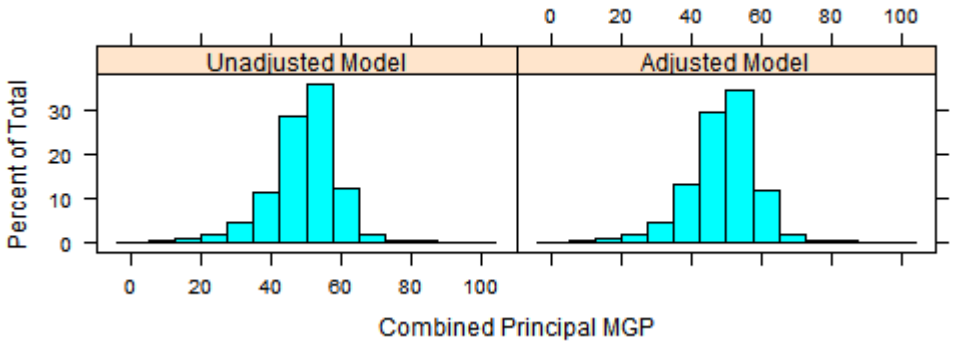


Figure 5 shows that for principals, the results are more lumped together in the center than for teachers, with a slight skew creating a longer "tail" on the left-hand side.

Figure 5: Distribution of Principal MGP by Model



Precision of the Mean Growth Percentiles

The caterpillar plot in Figure 6 is a random sample of 100 teacher MGPs taken from the New York data. The MGPs are sorted from lowest to highest with the corresponding 95 percent confidence interval showing the lower and upper limits of the MGP. Figure 7 shows the same for

principals where larger underlying samples mean that there is substantially less variation in the MGP and the error bars are narrower.

Figure 6: Overall MGP with 95% Confidence Interval Based on Random Sample of 100 Teachers

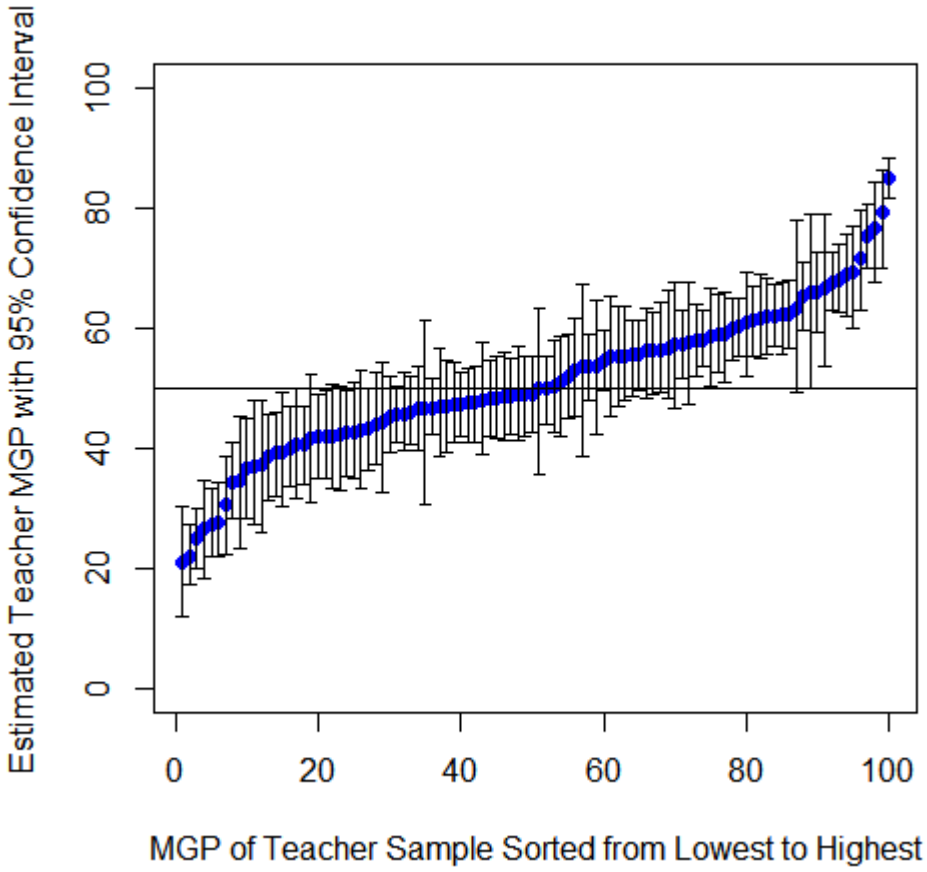


Figure 7: Overall MGP with 95% Confidence Interval Based on Random Sample of 100 Principals

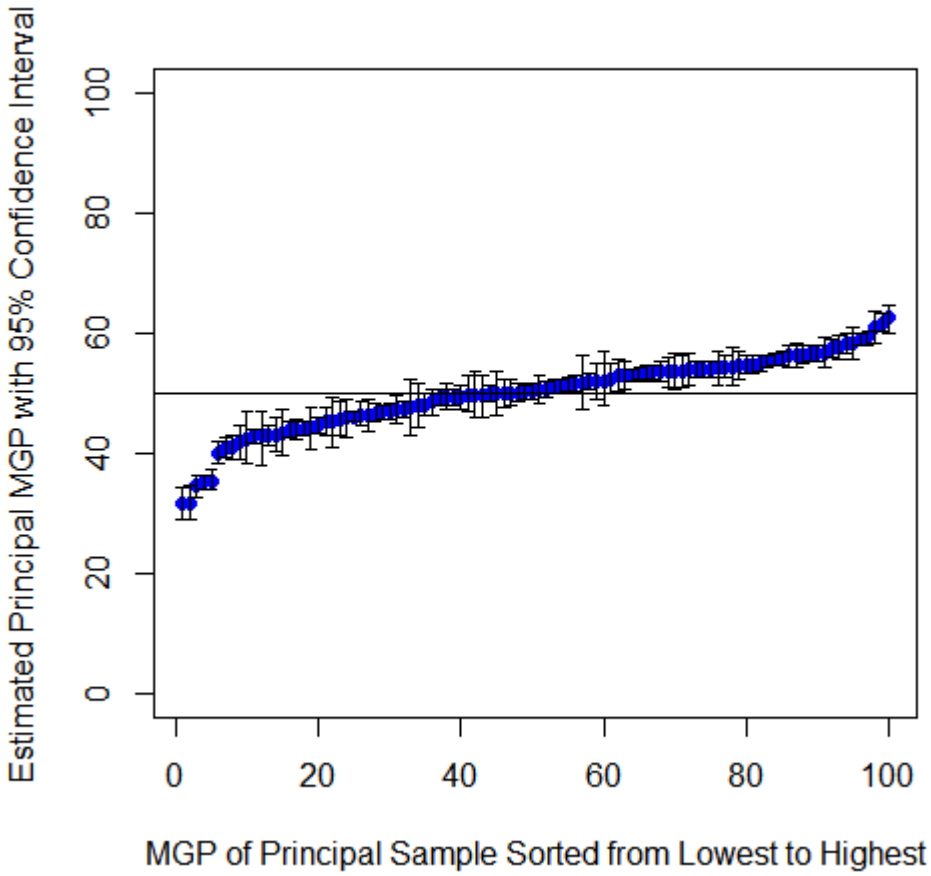


Table 4 provides the number of educators MGPs by model that are significantly below the mean for that educator type, using the 95 percent confidence intervals from the bootstrap. In all cases, the percent exceeding the mean is larger than what would be expected by chance alone (i.e., 5% expected by chance, or 2.5% for each table entry).

Table 4: Percent of Educator MGPs Above or Below Mean at the 95% Confidence Level

Educator Type	Model	Below Mean	Above Mean
Teacher	Adjusted	25.0%	27.2%
	Unadjusted	25.0%	26.9%
Principal	Adjusted	31.7%	37.9%
	Unadjusted	31.5%	38.4%

These figures provide a means to visually check the precision of the MGPs. However, it may also be useful to examine a statistic to gauge the precision of the teacher-level MGPs. The statistic, which we specify as ρ, is the ratio of the mean standard error to the standard deviation between teacher MGPs.

$$\rho = \frac{\bar{\sigma}}{sd(\hat{\theta}_j)}$$

Where $\bar{\sigma}$ is the mean standard error of the MGP and $sd(\hat{\theta}_j)$ is the standard deviation between teacher MGPs. In theory, the highest possible value is 1 and it represents a complete lack of precision in the measure -- when the ratio is 1, the variation in MGPs is explained entirely by sampling variation. Smaller values of ρ are associated with more precisely measured MGPs.

Table 5 below provides the mean standard errors, the standard deviation, and the value of ρ for the adjusted and unadjusted models by grade (again, for combined-subject MGPs). The values of the ratio (ρ) quantify imprecision in the estimates.

Table 5: Mean Standard Errors, Standard Deviation, and Value of ρ for Adjusted and Unadjusted Models by Grade (teachers only)

Grade (Teachers)	Unadjusted Mean SE	Adjusted Mean SE	Unadjusted Standard Deviation	Adjusted Standard Deviation	Unadjusted ρ	Adjusted ρ
4	4.19	4.42	13.77	13.34	0.30	0.33
5	4.71	4.72	12.83	12.67	0.37	0.37
6	4.74	4.77	12.65	12.38	0.37	0.39
7	4.5	4.52	12.04	11.88	0.37	0.38
8	4.61	4.67	13.15	12.98	0.35	0.36
Principals	1.92	1.93	7.56	7.60	0.25	0.25

Impact Data Results

Table 6 below provides the correlations of the combined-subject MGP (or for teachers with only one subject, their single-subject MGP) with five classroom characteristics: percent ELL, percent SWD, percent poverty (or economically disadvantaged (ED)), and the mean prior ELA or Mathematics score of the students.[1] This is done both with unadjusted and adjusted MGPs, with the adjusted model including demographic variables for individual student[2] The correlations

[1] As described earlier in this document, NYSED's regulations permit these three specific control variables at the individual student level for inclusion in the model, and all three were selected after discussion with New York's Task Force and other stakeholders.

[2] The impact of these demographic characteristics on the expected value of students' current test score used to compute SGPs can be seen through the model coefficients presented in Appendix H. The inclusion of these variables serves to make SGPs for students with different demographic characteristics comparable, given the prior test scores included in the model.

show that including these demographic variables in the model reduces but does not eliminate correlations between MGPs and classroom aggregates of model variables (and school-level aggregates for principals presented in Table 7).

The scatter plots shown in Figures 8 through 12 provide visual representations of the data underlying the correlations for teachers. Ideally the MGPs of teachers will be uncorrelated with classroom characteristics, given the student level variables included in the model. They may, however, be correlated for several reasons. One is that classroom factors may affect the learning environment – for example, a teacher in a highly-concentrated ELL class may offer different instruction to a student with a given test history than another teacher with few ELL students. The correlation may also affect differences in teachers and performance of students that vary by characteristics of classrooms and schools that are not fully captured by individual student prior test scores.

Figure 8 shows the relationship between MGPs and the classroom share of ELL students. The correlation is low, t, and indicates that the MGP is unrelated to percent ELL in a class. This suggests that, in general, there is no significant advantage or disadvantage in terms of MGP if teachers have a higher share of ELL students in their classes. The dispersion of points around the line does indicate that there is a great deal of variability in MGPs across shares of ELL students in a classroom – for this reason, some teachers appear to be advantaged by having ELL students and others are not.

Table 6 indicates that there is a moderate correlation with the percent SWD and percent economically disadvantaged in a class. This correlation is negative, indicating that as the percent SWD or percent economically disadvantaged in a class increases, the MGP tends to decrease.

The correlation with mean prior scores is also moderate in size and positive. This suggests that the MGP tends to be higher for classes serving students with higher mean levels of prior ability in both ELA and in Mathematics, even though SGPs for individual students have a low or moderate negative correlation with SGPs (Table 3) . The finding that MGPs for teachers are larger in classrooms where mean prior achievement is higher suggests that students with higher SGPs (given their prior test scores) may be assigned to teachers who have classrooms with larger shares of students who have higher test scores.

Table 6: Teacher MGP Correlated with Class Characteristics

Percent	Unadjusted Model	Adjusted Model
ELL in Class	-0.07	-0.03
SWD in Class	-0.26	-0.15
ED in Class	-0.18	-0.12
Mean Prior ELA	0.33	0.26
Mean Prior Math	0.31	0.28

Figure 8: Relationship of Teacher MGP Scores to Percent of ELL Students

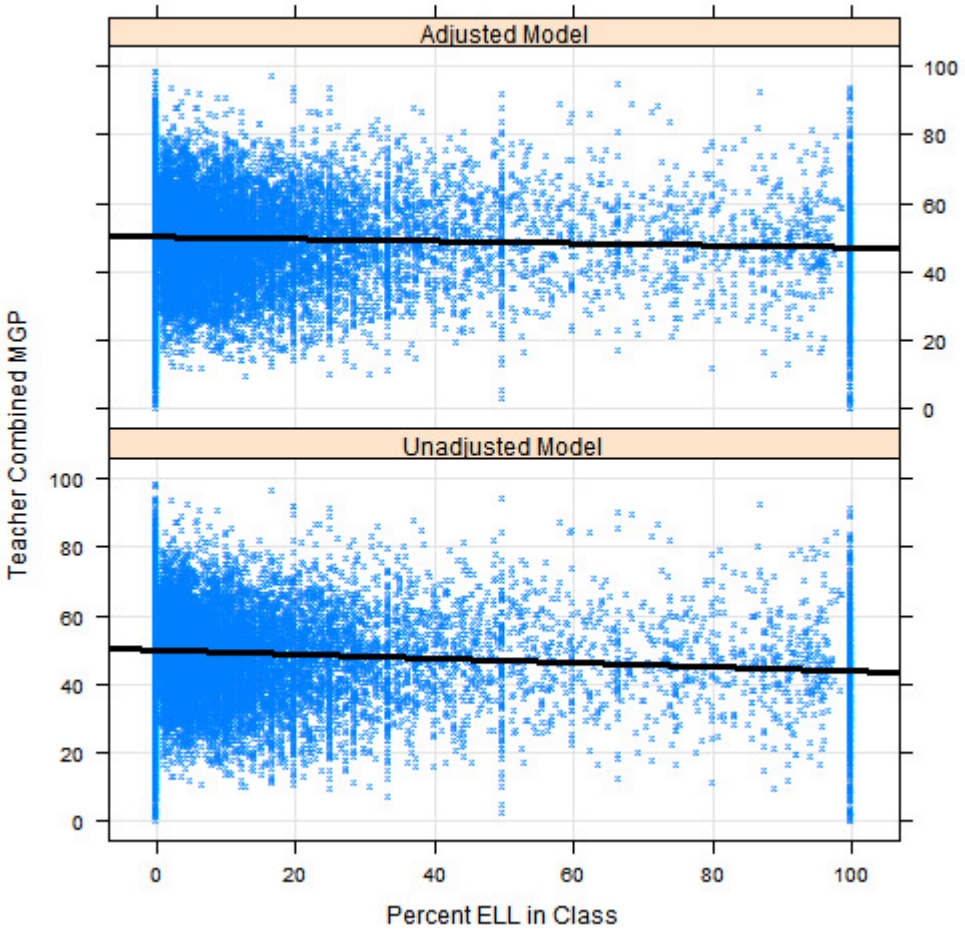


Figure 9: Relationship of Teacher MGP with Percent SWD in Class

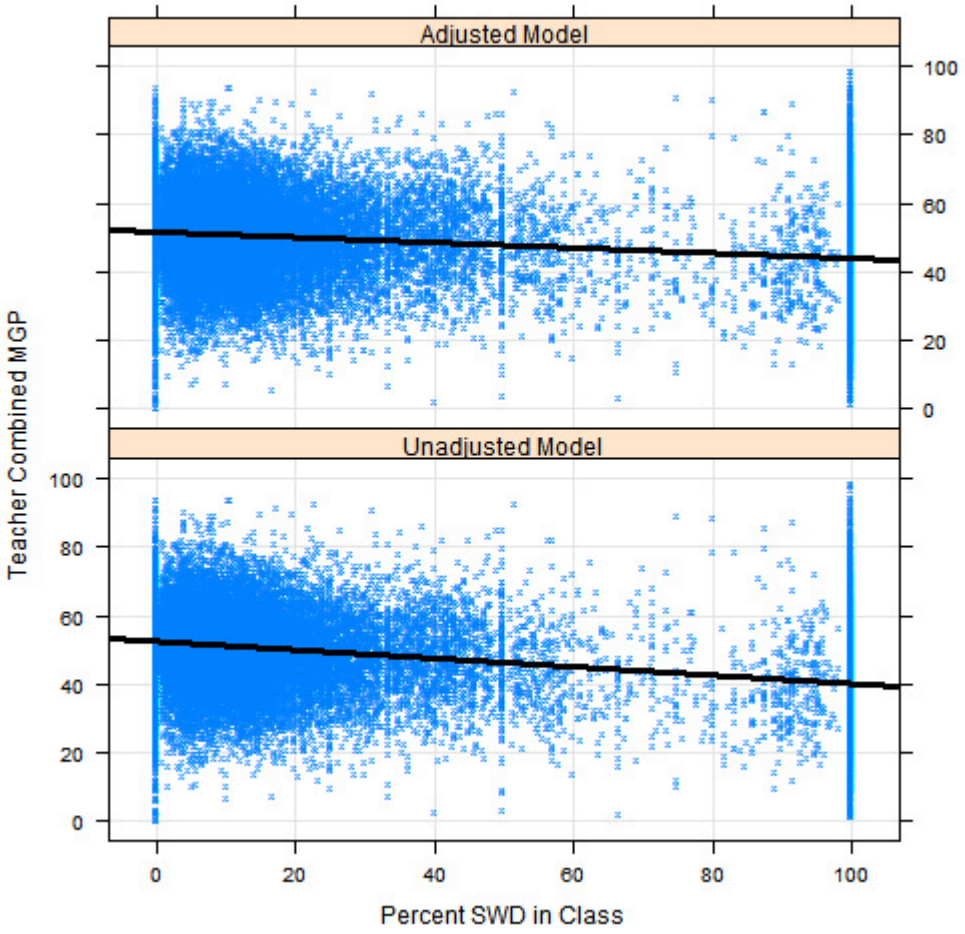


Figure 10: Relationship of Teacher MGP Scores to Percent of Economically Disadvantaged Students

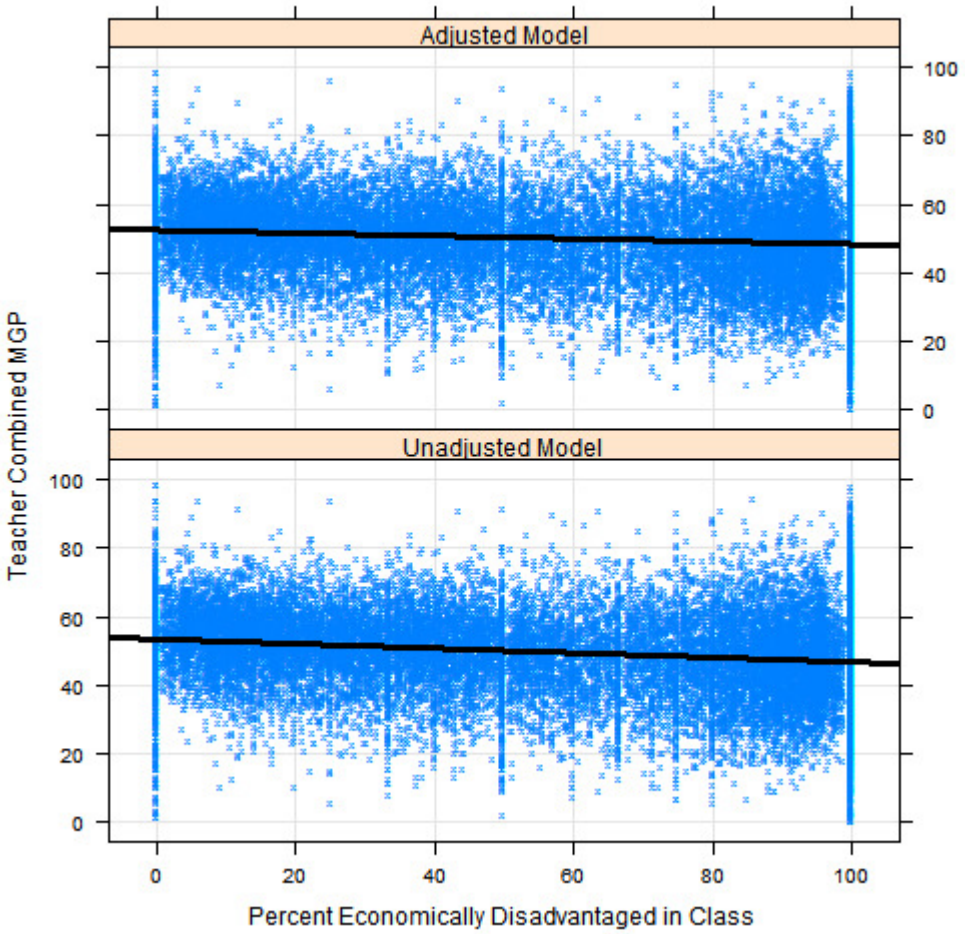
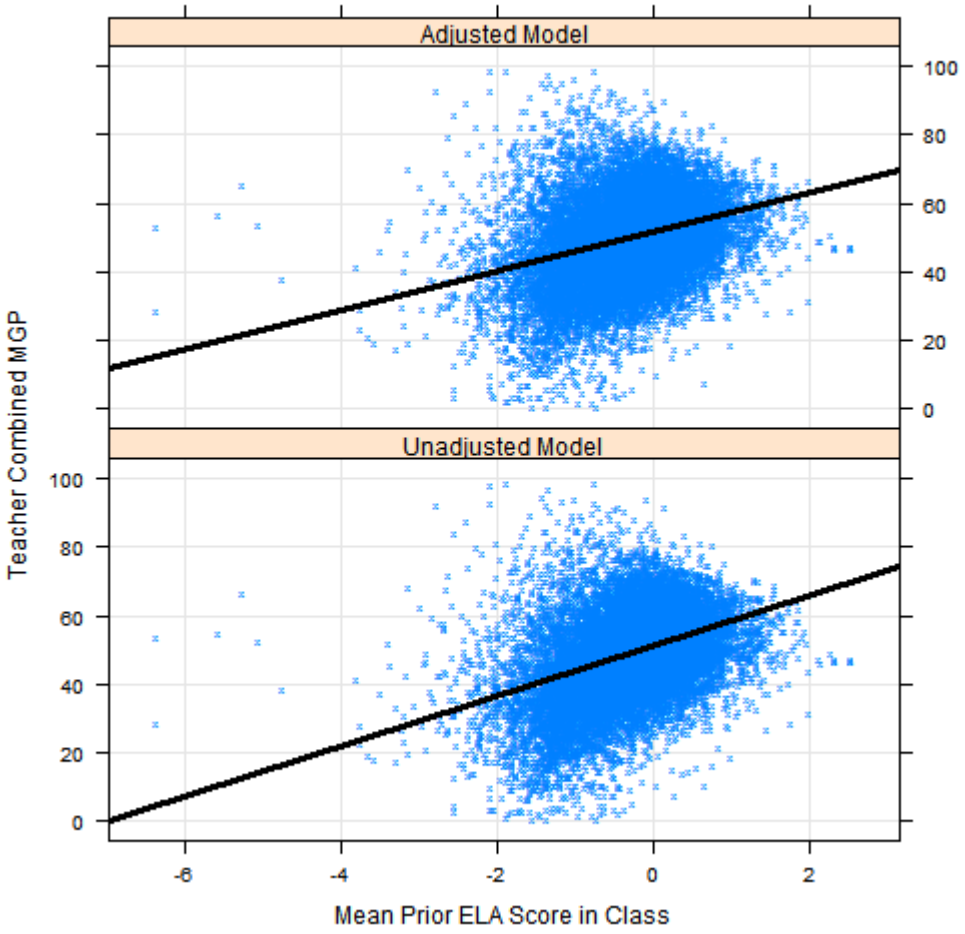
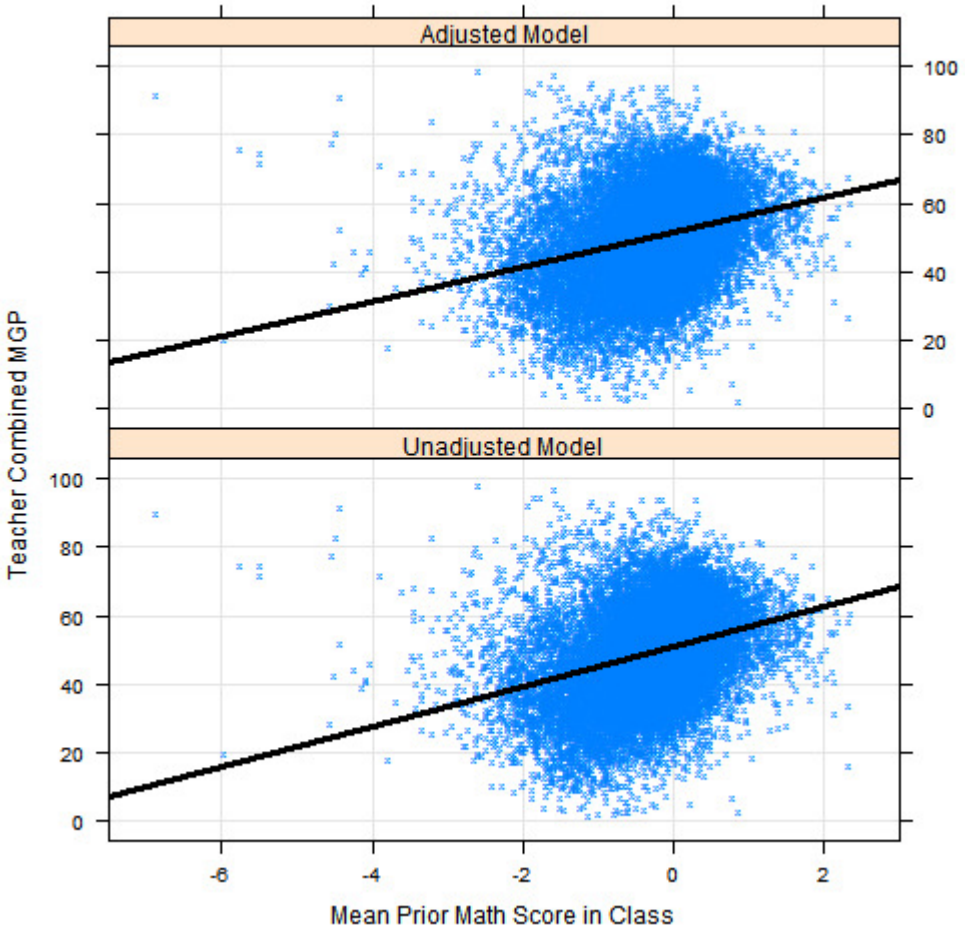


Figure 11: Relationship of Teacher MGP Scores to Mean Prior ELA Scores



Note: prior scores are Z-scores.

Figure 12: Relationship of Teacher MGP Scores to Mean Prior Mathematics Scores



Note: prior scores are Z-scores.

Table 7 below provides the observed correlations of principal MGPs with the five school characteristics matching those presented in table 6 for teachers: percent ELL in school, percent SWD in school, percent poverty in school, and the mean prior ELA or Mathematics score of the students. Because these are school-level characteristics, they are combined-subject, combined-grade results. The scatter plots shown in Figures 13 through 17 provide visual representations of these data. Results are presented for both unadjusted MGPs and adjusted MGPs, and as was the case at the teacher level, the inclusion of student-level covariates reduces correlations. .

The low correlation of the aggregate ELL variable to MGP s indicates that the MGP is unrelated to percent ELL in a school. This suggests that, in general, there is no significant advantage or disadvantage to principals conditional on this school characteristic, although there may be some relative advantage or disadvantage to some principals, given the variability in the data.

However, there is a moderate correlation with the percent SWD, and percent economically disadvantaged in a school -- in the unadjusted model only. This correlation is negative, indicating that as the percent SWD or percent economically disadvantaged in a school increases, the MGP tends to decrease.

The correlation with mean prior scores is also moderate in size and positive. This suggests that the MGP tends to be higher in schools with students with higher mean levels of ability in both ELA and in Mathematics.[3] The finding that MGPs for principals are larger in schools in which mean prior achievement is higher suggests that students with higher SGPs (given their prior test scores) may attend schools in which there are larger shares of students who have higher test scores.

Table 7: Principal MGP Correlated with Demographic Characteristics

Percent	Unadjusted Model	Adjusted Model
ELL in School	-0.05	0.02
SWD in School	-0.40	-0.29
ED in School	-0.19	-0.07
Mean Prior ELA	0.53	0.48
Mean Prior Math	0.58	0.54

[3] Appendix J provides additional impact analyses.

Figure 13: Relationship of Principal MGP Scores to Percent of ELL Students

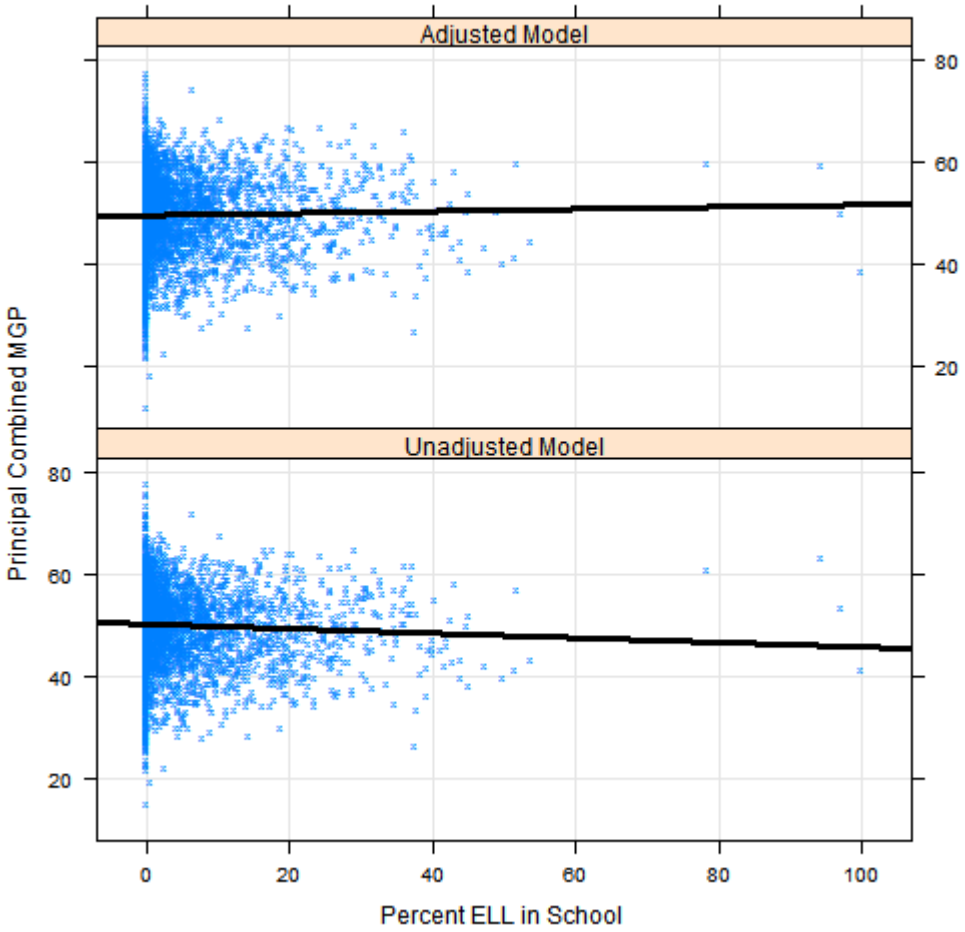


Figure 14: Relationship of Principal MGP with Percent SWD in School

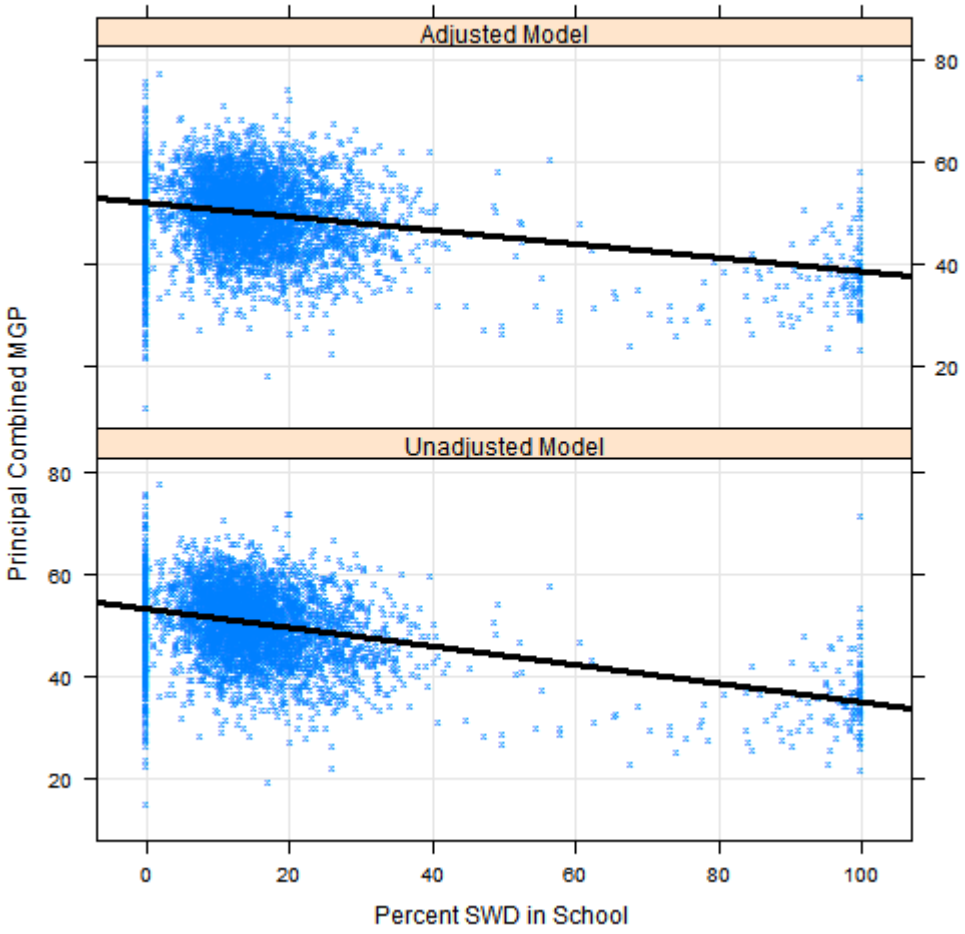


Figure 15: Relationship of Principal MGP Scores to Percent of Economically Disadvantaged Students

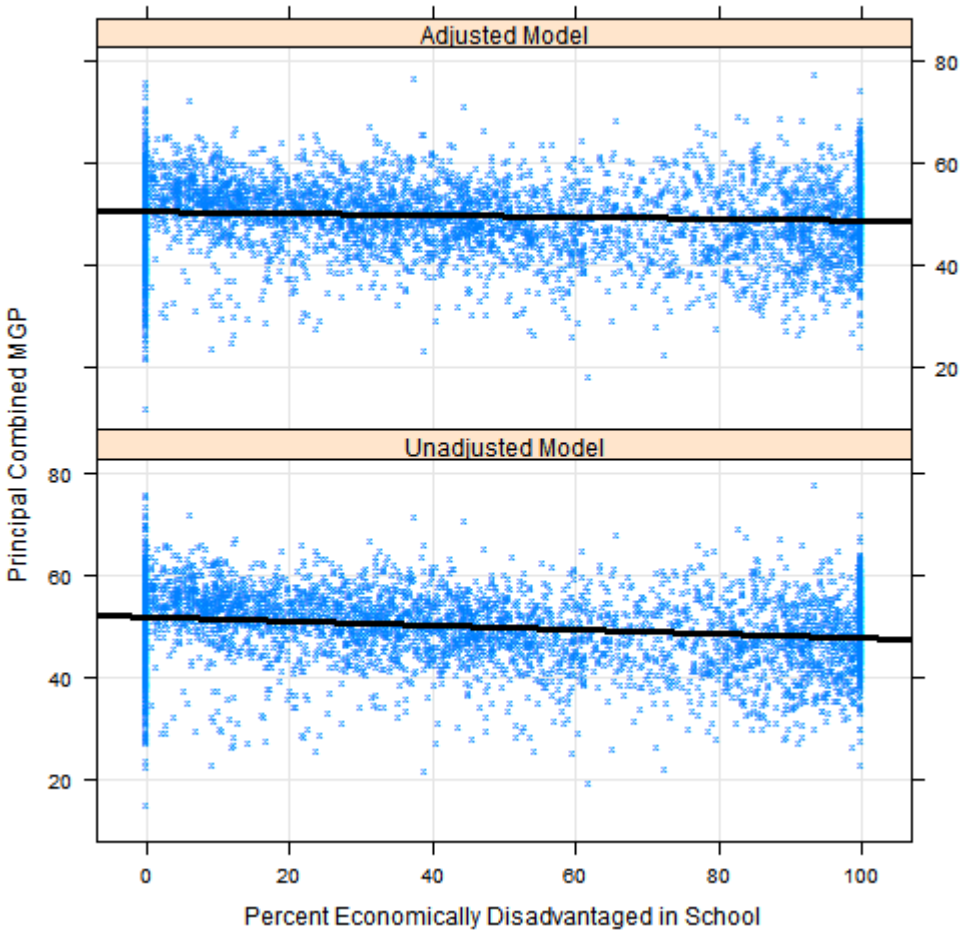
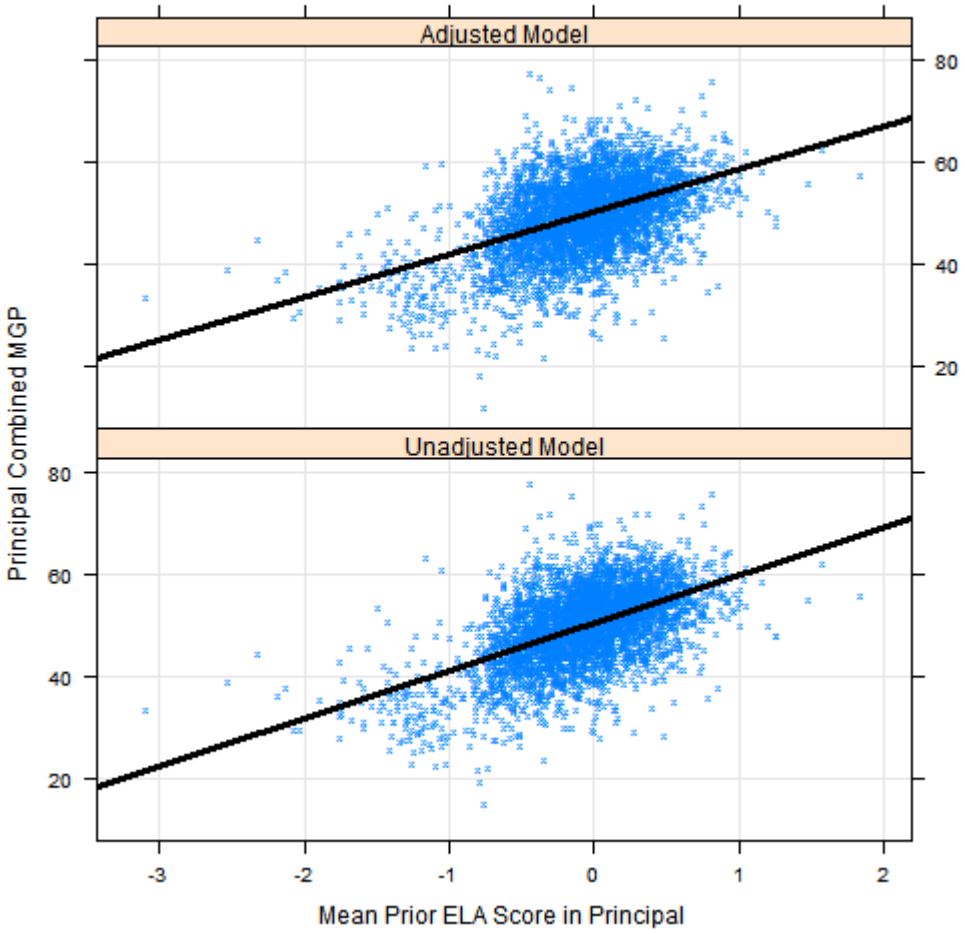
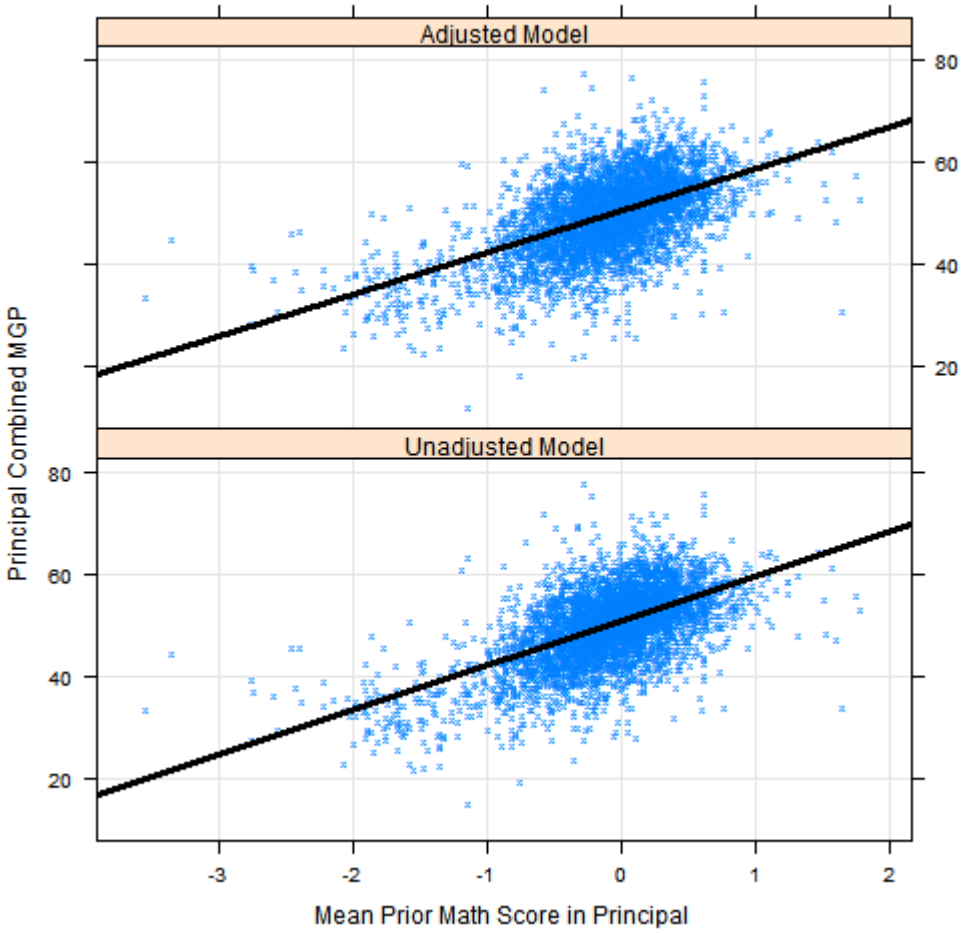


Figure 16: Relationship of Principal MGP Scores to Mean Prior ELA Scores



Note: prior scores are Z-scores.

Figure 17: Relationship of Principal MGP Scores to Mean Prior Mathematics Scores



Note: prior scores are Z-scores.

CONCLUSION

The model selected to estimate growth scores for New York State represents a first effort to produce fair and accurate estimates of individual teacher and principal effectiveness based on a limited set of data. The models were selected and developed based on technical and data considerations and on the recommendations of a variety of stakeholders. As new assessments are implemented, other data improvements are made, and the State pilots its new evaluation system, additional changes and adjustments to the growth model may be made or a transition to new models may be undertaken. This may involve including additional variables at the classroom and school levels to help adjust for differences in teacher and principal outcomes not captured by student-level variables.

REFERENCES

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practices 28*(4), 42–51.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7(1), 1–26.

Greene, W. (2001). Estimating econometric models with fixed effects. Leonard N. Stern School, Finance Department, Working Paper Series 01-10. New York: New York University, Leonard N. Stern School of Business.

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives 15*(4), 57–67.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation.

Wei, Y., & Carroll, R. J. (2009). Quantile regression with measurement error. *Journal of the American Statistical Association 104*, 1129–1143.

Appendix A. Task Force and Technical Advisory Committee Members

Participant	Affiliation
Technical Advisory Committee	
Dan Goldhaber	University of Washington
Hamp Lankford	State University of New York at Albany
Daniel F. McCaffrey	RAND
Jonah Rockoff	Columbia University
Tim R. Sass	Georgia State University
Douglas Staiger	Dartmouth College
Marty West	Harvard University
James A. Wyckoff	University of Virginia
Task Force	
Heather Adams	NYSUT
Jackie Bennett	UFT
Jack Bierwirth	NYSCOSS; Superintendent, Herricks Public Schools
Joanna Cannon	District Accountability Administrator, NYC Dept of Education
Nicole Catapano	SAANYS; Coordinator for Data Analysis, WSWHE Boces
Jim Dexter	District Superintendent, WSWHE Boces
Paul Gasparini	SAANYS; Principal Jamesville-Dewitt High School
Janella Hinds	UFT
Jon Hunter	NYSCOSS; Superintendent, Fairport Central School District
Maria Neira	NYSUT
Chris Ogno	CSA; Principal, P.S. 247 The NYC College Partnership School
Pilar Sokol	NYSSBA; Deputy General Counsel
Sue Winer	CSA; Instructional Literacy Coach, Syracuse City School District
Jay Worona	NYSSBA; General Counsel
Hamilton Lankford	Technical Advisor
Jonah Rockoff	Technical Advisor
James Wyckoff	Technical Advisor

Appendix B. Data Processing Rules and Results

Table B-1: Data Processing Rules (version 11 as of March 29, 2012)

Description		Rationale
D.1	Exclude from analysis students with missing, invalid, or duplicate ID numbers after matching procedures have been implemented	Not possible to merge student records over time
D.2	Exclude from analysis all students with no current achievement scores or standard error of measurement (SEM) for those scores in grades and subjects used in growth analysis	Not possible to generate growth scores without current achievement
D.3	Exclude from analysis all students with no prior achievement scores in grades and subjects used in growth analysis	Not possible to generate growth scores without any prior achievement
D.3b	Exclude from analysis all students with no immediate prior achievement scores or SEM for those scores in grades and subjects used in growth analysis	NYSED decision
D.4	Include in analysis all students with current achievement scores, and at least the immediate prior achievement score, and SEM for those scores	Possible to generate growth scores with partial prior achievement – helps maintain sample size and allows for modeling the impact of missing scores
D.5	Include in analysis students with missing demographic data on variables included in models	Possible to estimate models with missing data employing missing flags although too many missing flags will bias coefficients
D.5b	Include in analysis students with multiple disability codes (e.g., 385 – learning disability and 462 – other health impairment) but recode their disability code to be 473 – multiple disability	No reason to create a category for these few students or to consider them different from what other multiple disability conditions may be
D.5c	Include in analysis students with different values for ELL variables across records (e.g., Yes in one record, No in another) but set this variable to No	No way to tell which record is correct, so removing both records from the file with ELL subcategories file. Students who are not in the file with ELL subcategories are assumed not to be ELL and students who are no longer in the file with ELL subcategories are coded the same as other students not in that file.
D.6	Exclude from analysis students with an atypical grade progression, including students with multiple records in current or prior year(s) with contradictory grades (e.g. enrolled in one school as grade 5 and in another as grade	Not possible to validate which grade is correct.

Description		Rationale
	6; tested in one year as grade 5 and in the following year as grade 4).[4] Exception: retained students, see D.6b	
D.6b	Retained students – regardless of year retained - are assumed to have an atypical grade progression. Students retained in the most recent year are excluded according to D.6. However, include other students who have a test history with *some* of the appropriate grade progression. For example, if a grade 8 student has prior scores for grades 6 three years ago, 6 two years ago, and 7 last year, keep the second grade 6 score and the grade 7 score (setting the first grade 6 score to 0 and the missing indicator for this variable to 0). Similar rules apply to students who skip a grade	Retained students are not directly comparable to non-retained students in their grade. They are kept in the sample if the retention occurred in previous years to maintain sample size, and with the expectation that earlier scores will have little impact on the analysis.
D.7	Exclude from analysis all students with multiple records in current or prior year(s) with contradictory test scores[5]	Not possible to validate which score is correct
D.8	Exclude from analysis all students with only **invalidated**, or "did not attempt," or out-of-range test scores in current or prior years[6]	Valid test scores needed for value-added model
D.9	Exclude from analysis students flagged as "exclude from evaluation"	Other rules dictated by client/policy/legislation

[4] Note: if multiple years of prior scores used, exclude years with contradictory grades e.g., if two years of prior achievement used, and contradictory grades in only one year, can still include the other prior score/grade and keep student in analysis.
[5] *ibid*.
[6] Note: if multiple years of prior scores used, exclude years with invalidated or "did not attempt" or out-of-range scores e.g., if two years of prior achievement used, and these scores exist in only one year, can still include the other prior score and keep student in analysis.

Table B-2: ELA Data Processing

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
2007-08 Score File Processing				
Number of Records in the input ScoreFile	All	2007-08	2543074	-
Number of Records in the Score File after keeping only the records for the current subject	All	2007-08	1266669	1,276,405
Number of Records in the Score File after deleting blank/Invalid SSID	All	2007-08	1266669	0
Number of Records in the Score File after deleting Out of Range Scores	All	2007-08	1266046	623
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2007-08	1265994	52
Number of Records in the Score File after deleting all duplicates by SSID	All	2007-08	1265530	464
Number of Records in the Score File after deleting records with missing SEM	All	2007-08	1265529	1
2008-09 Score File Processing				
Number of Records in the input ScoreFile	All	2008-09	2531671	-
Number of Records in the Score File after keeping only the records for the current subject	All	2008-09	1260262	1,271,409
Number of Records in the Score File after deleting blank/Invalid SSID	All	2008-09	1260262	0
Number of Records in the Score File after deleting Out of Range Scores	All	2008-09	1259423	839
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2008-09	1259419	4
Number of Records in the Score File after deleting all duplicates by SSID	All	2008-09	1258869	550
Number of Records in the Score File after deleting records with missing SEM	All	2008-09	1258868	1
2009-10 Score File Processing				
Number of Records in the input ScoreFile	All	2009-10	2540947	-
Number of Records in the Score File after keeping only the records for the current subject	All	2009-10	1263223	1,277,724
Number of Records in the Score File after deleting blank/Invalid SSID	All	2009-10	1263223	0
Number of Records in the Score File after deleting Out of Range Scores	All	2009-10	1262133	1,090
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2009-10	1262131	2

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
Number of Records in the Score File after deleting all duplicates by SSID	All	2009-10	1261612	519
Number of Records in the Score File after deleting records with missing SEM	All	2009-10	1261612	0
2010-11 Score File Processing				
Number of Records in the input ScoreFile	All	2010-11	2540183	.
Number of Records in the Score File after keeping only the records for the current subject	All	2010-11	1263809	1,276,374
Number of Records in the Score File after deleting blank/Invalid SSID	All	2010-11	1263809	0
Number of Records in the Score File after deleting Out of Range Scores	All	2010-11	1262568	1,241
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2010-11	1262566	2
Number of Records in the Score File after deleting all duplicates by SSID	All	2010-11	1262034	532
Number of Records in the Score File after deleting records with missing SEM	All	2010-11	1262033	1
Number of Records in the Score File after deleting grade3 records	All	2010-11	1059894	-
Number of Records in the data for a specific grade	4	2010-11	215353	-
Number of Records in the data for a specific grade	5	2010-11	205727	-
Number of Records in the data for a specific grade	6	2010-11	213892	-
Number of Records in the data for a specific grade	7	2010-11	205177	-
Number of Records in the data for a specific grade	8	2010-11	219745	-
Merging 2010-11 Score File with Prior Year Score Files				
Number of Records in the Merged Score File after deleting records with prior grade greater than current grade	All	2010-11	1059894	-
Number of Records in the Merged Score File after deleting records without immediate prior score	All	2010-11	964804	95,090
Number of Records for a specific grade after merging 3 years of data	4	2010-11	190970	-
Number of Records for a specific grade after merging 3 years of data	5	2010-11	196103	-
Number of Records for a specific grade after merging 3 years of data	6	2010-11	191835	-
Number of Records for a specific grade after merging 3 years of data	7	2010-11	192642	-
Number of Records for a specific grade after merging 3 years of data	8	2010-11	193254	-
SGP File Processing				
Number of Records in the SGP file	All	2010-11	2398530	-
Number of Records in the SGP file after deleting Grade 3 records and keeping the ELA records only	All	2010-11	939991	1,458,539

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
Number of Records in the SGP file after deleting all duplicate SSID	All	2010-11	939991	-
Additional SWD and ELL Data File Processing				
Number of Records in the Subgroup file after deleting duplicates	All	2010-11	782647	157,344
Teacher-Course File Processing				
Number of Records in TeacherCourseFile	ALL	2010-11	1569394	-
Number of Records in TeacherCourseFile after subsetting for subject records	ALL	2010-11	721016	848,378
Number of Records in TeacherCourseFile after deleting duplicates	ALL	2010-11	719866	1,150
Merging Teacher-Course File to Merged Student Test Score Files*				
Number of Records in StudentTeacherCourseFile	ALL	2010-11	1017996	-
Number of Records in StudentTeacherCourseFile after keeping only 6 records per student	ALL	2010-11	1017993	3
Description of Final Reference File Used For Analysis*				
Number of Records in the Reference File	ALL	2010-11	964804	53,189
Number of Records in the Reference File for a specific grade	4	2010-11	190970	-
Number of Records in the Reference File for a specific grade	5	2010-11	196103	-
Number of Records in the Reference File for a specific grade	6	2010-11	191835	-
Number of Records in the Reference File for a specific grade	7	2010-11	192642	-
Number of Records in the Reference File for a specific grade	8	2010-11	193254	-

*Note that the number of observations shown in the merging teacher-course and student test score file represents the number of unique teacher-course-student links, while the number of observations in the final reference file description represents the number of students. That is, the difference is due to the fact that a student may be linked to multiple teachers/courses. The file is transformed for statistical analysis to include one record per student.

Table B-3: Math Data Processing

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
2007-08 Score File Processing				
Number of Records in the input ScoreFile	All	2007-08	2543074	-
Number of Records in the Score File After keeping only the records for the current subject	All	2007-08	1276405	1,266,669
Number of Records in the Score File after deleting blank/Invalid SSID	All	2007-08	1276405	0
Number of Records in the Score File after deleting Out of Range Scores	All	2007-08	1275124	1,281
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2007-08	1275122	2
Number of Records in the Score File after deleting all duplicates by SSID	All	2007-08	1274664	458
Number of Records in the Score File after deleting records with missing SEM	All	2007-08	1274661	3
2008-09 Score File Processing				
Number of Records in the input ScoreFile	All	2008-09	2531671	-
Number of Records in the Score File After keeping only the records for the current subject	All	2008-09	1271409	1,260,262
Number of Records in the Score File after deleting blank/Invalid SSID	All	2008-09	1271409	0
Number of Records in the Score File after deleting Out of Range Scores	All	2008-09	1270385	1,024
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2008-09	1270385	0
Number of Records in the Score File after deleting all duplicates by SSID	All	2008-09	1269875	510
Number of Records in the Score File after deleting records with missing SEM	All	2008-09	1269875	0
2009-10 Score File Processing				
Number of Records in the input ScoreFile	All	2009-10	2540947	-
Number of Records in the Score File After keeping only the records for the current subject	All	2009-10	1277724	1,263,223
3:Number of Records in the Score File after deleting blank/Invalid SSID	All	2009-10	1277724	0
Number of Records in the Score File after deleting Out of Range Scores	All	2009-10	1276339	1,385
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2009-10	1276336	3
Number of Records in the Score File after deleting all duplicates by SSID	All	2009-10	1275807	529

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
Number of Records in the Score File after deleting records with missing SEM	All	2009-10	1275807	0
2010-11 Score File Processing				
Number of Records in the input ScoreFile	All	2010-11	2540183	-
Number of Records in the Score File After keeping only the records for the current subject	All	2010-11	1276374	1,263,809
Number of Records in the Score File after deleting blank/Invalid SSID	All	2010-11	1276374	0
Number of Records in the Score File after deleting Out of Range Scores	All	2010-11	1275437	937
Number of Records in the Score File after deleting one of the duplicates by SSID Item_Description ScaleScore School_ID District_ID S_EnrolledGrade S_Lastname S_FirstName	All	2010-11	1275435	2
Number of Records in the Score File after deleting all duplicates by SSID	All	2010-11	1274887	548
Number of Records in the Score File after deleting records with missing SEM	All	2010-11	1274886	1
Number of Records in the Score File after deleting grade3 records	All	2010-11	1070500	-
Number of Records in the data for a specific grade	4	2010-11	217507	-
Number of Records in the data for a specific grade	5	2010-11	208165	-
Number of Records in the data for a specific grade	6	2010-11	216033	-
Number of Records in the data for a specific grade	7	2010-11	207340	-
Number of Records in the data for a specific grade	8	2010-11	221455	-
Merging 2010-11 Score File with Prior Year Score Files				
Number of Records in the Merged Score File after deleting records with prior grade greater than current grade	All	2010-11	1070500	-
Number of Records in the Merged Score File after deleting records without immediated prior score	All	2010-11	975200	95,300
Number of Records in the for a specific grade after merging 3 years of data	4	2010-11	193037	-
Number of Records in the for a specific grade after merging 3 years of data	5	2010-11	198196	-
Number of Records in the for a specific grade after merging 3 years of data	6	2010-11	193852	-
Number of Records in the for a specific grade after merging 3 years of data	7	2010-11	194896	-
Number of Records in the for a specific grade after merging 3 years of data	8	2010-11	195219	-
SGP File Processing				
Number of Records in the SGP file	All	2010-11	2398530	-
Number of Records in the SGP file after deleting Grade 3 records and keeping the Math records only	All	2010-11	943541	1,454,989

Data Processing Description	Grade	Year	Resulting # Obs After Exclusion	# Records excluded
Number of Records in the SGP file after deleting all duplicate SSID	All	2010-11	943541	0
Additional SWD and ELL Data File Processing				
Number of Records in the Subgroup file after deleting duplicates	All	2010-11	782647	160,894
Teacher-Course File Processing				
Number of Records in TeacherCourseFile	ALL	2010-11	1569394	-
Number of Records in TeacherCourseFile after subsetting for subject records	ALL	2010-11	658872	910,522
Number of Records in TeacherCourseFile after Deleting records that need to be excluded.	ALL	2010-11	658872	-
Number of Records in TeacherCourseFile after Deleting Duplicates	ALL	2010-11	658213	659
Merging Teacher-Course to Merged Student Test Score Files*				
Number of Records in StudentTeacherCourseFile	ALL	2010-11	1000619	-
Number of Records in StudentTeacherCourseFile after keeping only 6 records per student	ALL	2010-11	1000619	-
Description of Final Reference File Used for Analysis*				
Number of Records in the Reference File	ALL	2010-11	975200	25,419
Number of Records in the Reference File for a specific grade	4	2010-11	193037	-
Number of Records in the Reference File for a specific grade	5	2010-11	198196	-
Number of Records in the Reference File for a specific grade	6	2010-11	193852	-
Number of Records in the Reference File for a specific grade	7	2010-11	194896	-
Number of Records in the Reference File for a specific grade	8	2010-11	195219	-

*Note that the number of observations shown in the merging teacher-course and student test score file represents the number of unique teacher-course-student links, while the number of observations in the final reference file description represents the number of students. That is, the difference is due to the fact that a student may be linked to multiple teachers/courses. The file is transformed for statistical analysis to include one record per student.

Appendix C. Course Codes Used in Analysis

Table C-1: Relevant Course Codes

(State) Course Code	Course Name
1300	Grade 3 English Language Arts
2300	Grade 3 Mathematics
1400	Grade 4 English Language Arts
2400	Grade 4 Mathematics
1500	Grade 5 English Language Arts
2500	Grade 5 Mathematics
1600	Grade 6 English Language Arts
2600	Grade 6 Mathematics
1700	Grade 7 English Language Arts
2700	Grade 7 Mathematics
1800	Grade 8 English Language Arts
2800	Grade 8 Mathematics

Appendix D. Attribution Rules

Table D-1: Attribution Rules

Attribution Rule		Example	Notes
A.1	Students enrolled in the same relevant course in multiple time periods with the same or different teachers are attributed to those teachers for the proportion of time spent with each teacher prior to testing, as described by linkage and course duration and adjustment to linkage duration data	A grade 4 student begins in one school, transfers to another school for several months, then returns to the initial school where s/he is placed with the same initial grade 4 teacher. The teacher is responsible for the student for the proportion of course duration prior to testing that the student was in his/her classroom	Each relevant course is associated with one teacher only (last one reported) and the proportion of time is set at 100% (no additional data reported)
A.2	Students enrolled in multiple relevant courses with multiple teachers are attributed to each teacher for each course based on linkage and course duration and adjustment to linkage duration data	A grade 7 student is enrolled in three courses that are deemed relevant to the ELA assessment. The student takes one course with Teacher A, and the other two courses with Teacher B. The student will be attributed to both Teachers A and B with a weight of 1 for each teacher, taking into account linkage and course duration and adjustment to linkage duration data	Students enrolled in multiple relevant courses with multiple teachers are attributed to the single teacher associated with each course
A.3	Students not attributable to a teacher (i.e., not enrolled in a course code assigned to a teacher in a school) but with outcome test scores will be attributed to the school. Students not attributable to a school (i.e., students not receiving services at a school) but with outcome test scores will be attributed to the district.	A grade 4 student is not enrolled in the identified relevant grade 4 ELA or Math course, but has taken the State test.	

Attribution Rule		Example	Notes
A.4	Students in classrooms where no permanent teacher is in place (or is in not in place until after testing) are attributed to the school only.	A grade 5 classroom's permanent teacher resigns one month into the school year and is replaced by a long-term substitute for the rest of the school year. No individual teacher receives a value-added estimate, but these students are included in the school-level value-added estimate. (Rationale: long-term subs do not currently need to be reported for evaluation purposes as certified teachers are and we are unlikely to be able to produce value-added scores over time for these subs. However, schools may be responsible for finding appropriate permanent instructors for their students in which case they should be held responsible for student learning growth)	
A.5	Students in classrooms where the permanent teacher has been in place for any amount of time in the school year prior to testing are attributed to the teacher **proportional to the time they were taught by the teacher** as described by linkage and course duration and adjustment to linkage duration data	A grade 5 classroom has a substitute teacher for one month. A permanent teacher is then identified and teaches the students for the rest of the school year. This permanent teacher receives an estimate including the students in this classroom proportional to the time s/he taught in the classroom as described by linkage and course duration and adjustment to linkage duration data	

Attribution Rule		Example	Notes
A.6	Students tested in a school but enrolled in another and linked to a teacher in that other school are attributed to the latter school, where a student-teacher link exists – they are not attributed to school A	A student tests in school A but is enrolled in school B with teacher Y. This student is attributed to school B and teacher Y.	

Appendix E. Demographic Variable Definitions

For the beta model, New York has provided a Y/N variables to identify students as SWD.

NYSED has also provided ELL data with codes for different ELL statuses. These codes were provided for students who have ELL status, not for all students. New York has made the decision to model these data as simply Y/N. To accomplish this, AIR created single variables ELL coded as 1 if a student has any of the codes listed in the table below and 0 otherwise. Where students in the ELL file have duplicate, conflicting records (e.g., one record has ELL = Y and one has ELL = N), the data are considered invalid and deleted from the ELL file. When merged with the general data, these students are assumed to be non-ELL, similar to students not appearing in the ELL file.

ELL Code	ELL Description
5709	English as a Second Language
5676	Bilingual Program
5687	Two-way Bilingual Education Program

The poverty indicator is a 0/1 variable that was provided as is by NYSED. It is used to determine which cohort members should be included in the economically disadvantaged group for district and school accountability. An economically disadvantaged student is a student who participates in, or whose family participates in, economic assistance programs such as:

- The Free- or Reduced-price Lunch Programs (Note that the United States Department of Agriculture has authorized the use of enrollment in free- and reduced-price lunch programs to identify students from low-income families for Title I reporting purposes.) Please consult the NYSED's Office of Child Nutrition Program Administration for guidelines;
- Social Security Insurance (SSI);
- Food Stamps;
- Foster Care;
- Refugee Assistance (cash or medical assistance);
- Earned Income Tax Credit (EITC);
- Home Energy Assistance Program (HEAP);
- Safety Net Assistance (SNA);
- Bureau of Indian Affairs (BIA); or
- Family Assistance: Temporary Assistance for Needy Families (TANF).

If one student in a family is identified as low income, all students from that household (economic unit) may be identified as low income.

Appendix F. Descriptive Results

This appendix provides high-level descriptive statistics for unadjusted MGPs, adjusted MGPs, and adjustment factors (the difference between unadjusted and adjusted MGPs). Provided first are overall descriptive statistics at the educator level aggregated across grades by subject (ELA, Math, and combined). Following that, we present the data for the grade/subject-specific MGPs.

Table F-1 shows the mean, standard deviation, minimum, and maximum for the unadjusted MGPs, adjusted MGPs, and adjustment factors for ELA, Math, and combined subjects. These data are taken from the teacher-level aggregates (across grades for ELA, Math, and combined; and across subjects for combined), without applying the minimum class size filter for reporting. The average MGP tends to be about 50 (as expected), with unadjusted MGPs slightly lower than adjusted MGPs on average. The adjustment factor for teachers tends to be positive (although only slightly). For combined subject MGPs, adjustment factors ranged from –14 to 21.

Table F-1: Descriptive Statistics for Teacher-Level MGPs and Adjustment Factors

Subject	Statistic	Mean	Standard Deviation	Minimum	Maximum	*N*
ELA	Unadjusted MGP	49.4	14.6	1	99	18,235
	Adjusted MGP	50.1	14.6	1	99	18,235
	Adjustment Factor	0.7	2.8	–11	21	18,235
Math	Unadjusted MGP	48.7	18.7	1	99	16,765
	Adjusted MGP	49.5	18.6	1	99	16,765
	Adjustment Factor	0.7	3.3	–14	22	16,765
Combined	Unadjusted MGP	49.0	15.4	1	99	24,735
	Adjusted MGP	49.6	15.4	1	99	24,735
	Adjustment Factor	0.6	2.7	–14	21	24,735

Table F-2 provides the same statistics for principals, combined subjects only. Here the unadjusted are slightly higher than the adjusted, on average.

Table F-2: Descriptive Statistics for Principal-Level MGPs and Adjustment Factors

Subject	Statistic	Mean	Standard Deviation	Minimum	Maximum	*N*
Combined	Unadjusted MGP	49.1	10.4	1	99	4,593
	Adjusted MGP	48.7	10.5	1	99	4,593
	Adjustment Factor	-0.46	1.9	-13.5	17	4,593

Table F-3 provides the mean, standard deviation, minimum, and maximum for unadjusted MGPs by grade and subject. The unadjusted MGPs have averages just under 50 with one exception (ELA grade 5).

Table F-3: Descriptive Statistics for Unadjusted MGPs by Grade and Subject

Subject	Grade	Mean	Standard Deviation	Minimum	Maximum	*N*
ELA	4	47.6	13.9	1	99	4,750
	5	52.1	18.8	1	99	4,753
	6	47.5	16.4	1	99	4,325
	7	48.8	13.4	1	97	3,544
	8	48.3	13.1	1	95	3,402
Math	4	46.7	18.8	1	98	4,737
	5	48.6	20.1	1	99	4,564
	6	47.9	20.7	1	99	3,403

	7	48.4	19.7	1	99	3,278
	8	49.2	17.8	5	99	3,133

Table F-4 provides the mean, standard deviation, minimum, and maximum for adjusted MGPs by grade and subject. The adjusted MGPs represent the teacher scores from the model which includes student covariates for poverty, ELL, and SWD status. The adjusted MGPs have averages just under 50 with one exception (ELA grade 5). In general, the adjusted MGPs tend to be slightly higher than the unadjusted MGPs seen in Table F-2.

Table F-4: Descriptive Statistics for Adjusted MGPs by Grade and Subject

Subject	Grade	Mean	Standard Deviation	Minimum	Maximum	*N*
ELA	4	49.2	14.4	1	99	4,750
	5	52.2	18.7	1	99	4,753
	6	48.5	15.9	1	99	4,325
	7	49.3	13.5	1	97	3,544
	8	49.1	13.2	2	97	3,402
Math	4	49.2	18.8	1	99	4,737
	5	49.3	20.0	1	99	4,564
	6	48.2	20.6	1	99	3,403
	7	48.4	19.6	1	99	3,278
	8	49.0	17.7	5	99	3,133

Table F-5 provides the mean, standard deviation, minimum, and maximum for the adjustment factors by grade and subject. In general, the average adjustment factor is slightly greater than 0, except for grade 8 Math. Grade 4 Math had the largest average adjustment of 2.5.

Table F-5: Descriptive Statistics for Adjustment Factors by Grade and Subject

Subject	Grade	Mean	Standard Deviation	Minimum	Maximum	*N*
ELA	4	1.6	3.9	–11	18	4,750
	5	0.0	1.4	–6	4	4,753
	6	1.1	3.6	–9	21	4,325
	7	0.5	1.6	–4.5	6	3,544
	8	0.9	2.8	–7	13	3,402
Math	4	2.5	5.5	–13	22	4,737
	5	0.8	2.2	–5	8	4,564
	6	0.3	2.0	–8	7	3,403
	7	0.0	1.5	–10	3	3,278
	8	–0.2	2.1	–14	3	3,133

Appendix G. Model Variance Components

For each grade and subject, the growth models were fit to the data with both teacher and school random effects. The model decomposes total variation in the outcome into three unique components: variance between teachers within a school, variance between schools, and variance between students within a class. Table G-1 provides the standard deviation of the components for all subjects, grades, and models. Figures G-1 through G-4 provide a graphical look at these same data, with bar charts corresponding to a subject and model combination, shown across grades 4 through 8.

Table G-1. Magnitude of Teacher and School Effect Standard Deviations

Subject	Grade	Model	Root Mean Square Error	Teacher	School
ELA	4	Unadjusted	20.89	2.37	4.01
		Adjusted	18.91	2.72	4.01
	5	Unadjusted	5.38	2.26	2.33
		Adjusted	5.38	2.25	2.31
	6	Unadjusted	6.95	1.77	2.17
		Adjusted	6.88	1.63	2.03
	7	Unadjusted	8.76	1.09	1.96
		Adjusted	8.59	1.11	1.95
	8	Unadjusted	11.22	1.19	1.95
		Adjusted	10.73	1.14	1.97
Math	4	Unadjusted	18.85	6.40	6.09
		Adjusted	17.52	5.97	5.97
	5	Unadjusted	9.87	5.40	4.05
		Adjusted	9.75	5.38	4.01
	6	Unadjusted	8.54	6.23	4.84
		Adjusted	8.50	6.13	4.82
	7	Unadjusted	3.98	4.96	3.77
		Adjusted	3.93	4.93	3.72
	8	Unadjusted	10.66	4.66	3.92
		Adjusted	10.61	4.57	3.84

Figure G-1. Magnitude of Teacher and School Effects—Unadjusted ELA Models by Grade

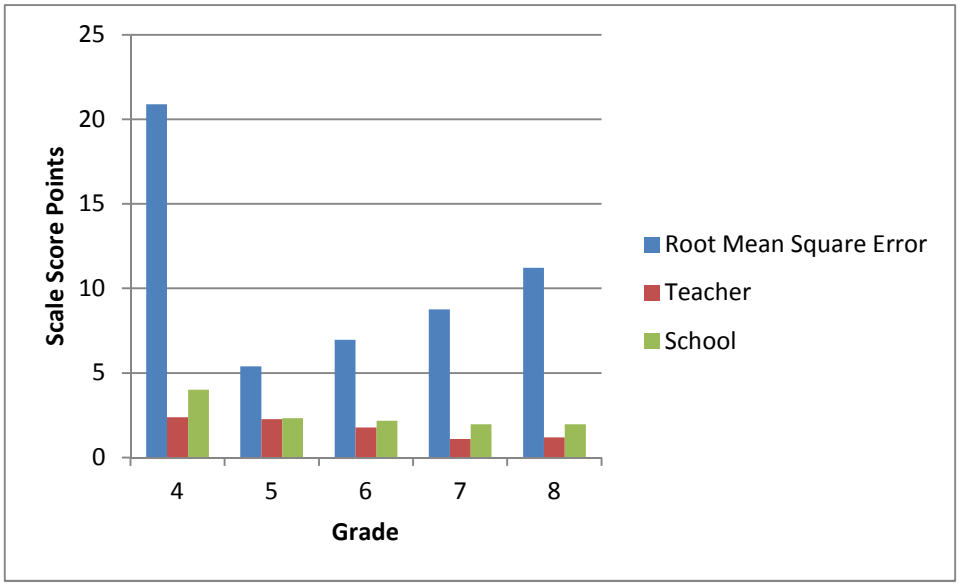


Figure G-2. Magnitude of Teacher and School Effects—Adjusted ELA Models by Grade

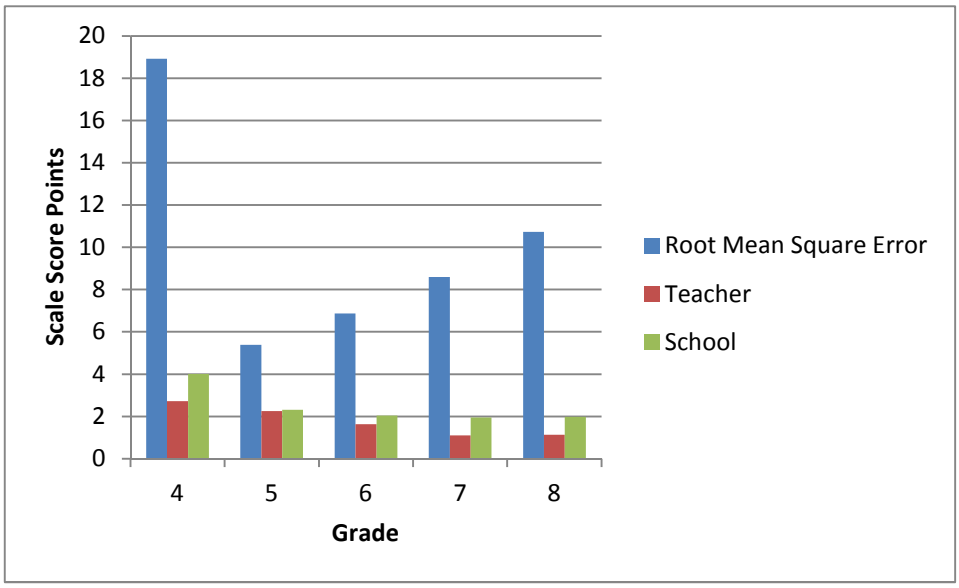


Figure G-3. Magnitude of Teacher and School Effects—Unadjusted Math Models by Grade

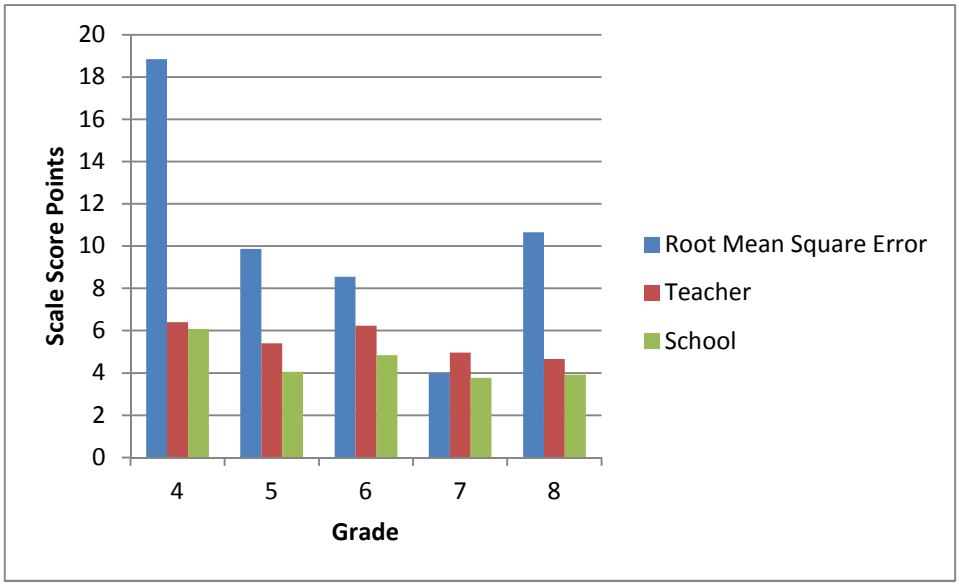
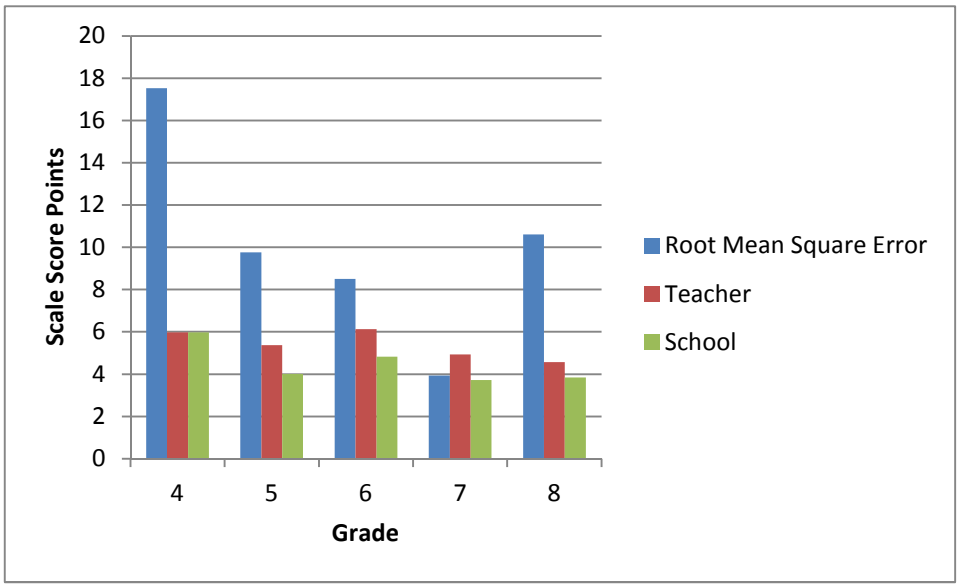


Figure G-4. Magnitude of Teacher and School Effects—Adjusted Math Models by Grade



Appendix H. Model Coefficients

Table H-1. Grade 4 ELA Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	–33.165	3.141	0.000
2009–10 Assessment Score	1.063	0.005	0.000

Table H-2. Grade 4 ELA Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	48.161	3.522	0.000
Student Poverty status	–2.758	0.143	0.000
Student SWD status	–7.471	0.176	0.000
Student ELL status	–2.350	0.233	0.000
2009-10 Assessment Score	0.945	0.005	0.000

Table H-3. Grade 5 ELA Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	282.398	0.861	0.000
2008–09 Prior Score Missing	91.927	2.655	0.000
2009–10 Assessment Score	0.436	0.004	0.000
2008–09 Assessment Score	0.138	0.004	0.000

Table H-4. Grade 5 ELA Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	286.306	1.061	0.000
2008–09 Prior Score Missing	88.549	2.735	0.000
Student Poverty status	–0.000	0.058	0.995
Student SWD status	–0.837	0.072	0.000
Student ELL status	0.894	0.104	0.000
2009–10 Assessment Score	0.435	0.004	0.000
2008–09 Assessment Score	0.133	0.004	0.000

Table H-5. Grade 6 ELA Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	328.898	0.941	0.000
2007–08 Prior Score Missing	49.765	1.424	0.000
2008–09 Prior Score Missing	146.145	1.700	0.000
2009–10 Assessment Score	0.199	0.002	0.000
2008–09 Assessment Score	0.225	0.003	0.000
2007–08 Assessment Score	0.074	0.002	0.000

Table H-6. Grade 6 ELA Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	348.757	1.022	0.000
2007–08 Prior Score Missing	44.872	1.398	0.000
2008–09 Prior Score Missing	135.191	1.686	0.000
Student Poverty status	–1.035	0.056	0.000
Student SWD status	–2.550	0.069	0.000
Student ELL status	–2.608	0.116	0.000
2009–10 Assessment Score	0.195	0.002	0.000
2008–09 Assessment Score	0.208	0.003	0.000
2007–08 Assessment Score	0.067	0.002	0.000

Table H-7. Grade 7 ELA Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	156.764	1.936	0.000
2007–08 Prior Score Missing	64.988	1.374	0.000
2008–09 Prior Score Missing	50.929	1.684	0.000
2009–10 Assessment Score	0.588	0.005	0.000
2008–09 Assessment Score	0.076	0.003	0.000
2007–08 Assessment Score	0.099	0.002	0.000

Table H-8. Grade 7 ELA Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	170.719	2.067	0.000
2007–08 Prior Score Missing	61.282	1.378	0.000
2008–09 Prior Score Missing	51.057	1.663	0.000
Student Poverty status	–0.776	0.066	0.000
Student SWD status	–1.261	0.081	0.000
Student ELL status	–0.211	0.147	0.150
2009–10 Assessment Score	0.573	0.005	0.000
2008–09 Assessment Score	0.076	0.002	0.000
2007–08 Assessment Score	0.094	0.002	0.000

Table H-9. Grade 8 ELA Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	86.237	1.547	0.000
2007–08 Prior Score Missing	90.994	2.284	0.000
2008–09 Prior Score Missing	78.217	2.307	0.000
2009–10 Assessment Score	0.597	0.005	0.000
2008–09 Assessment Score	0.120	0.004	0.000
2007–08 Assessment Score	0.138	0.003	0.000

Table H-10. Grade 8 ELA Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	117.889	1.763	0.000
2007–08 Prior Score Missing	81.903	2.212	0.000
2008–09 Prior Score Missing	76.673	2.226	0.000
Student Poverty Status	–0.759	0.079	0.000
Student SWD Status	–3.334	0.098	0.000
Student ELL status	–1.334	0.181	0.000
2009–10 Assessment Score	0.567	0.005	0.000
2008–09 Assessment Score	0.118	0.003	0.000
2007–08 Assessment Score	0.124	0.003	0.000

Table H-11. Grade 4 Math Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	–41.293	2.544	0.000
2009–10 Assessment Score	1.058	0.004	0.000

Table H-12. Grade 4 Math Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	34.901	2.612	0.000
Student Poverty Status	–3.906	0.138	0.000
Student SWD Status	–10.123	0.153	0.000
Student ELL Status	–2.334	0.213	0.000
2009–10 Assessment Score	0.952	0.004	0.000

Table H-13. Grade 5 Math Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	166.727	1.112	0.000
2008–09 Prior Score Missing	76.225	1.790	0.000
2009–10 Assessment Score	0.647	0.003	0.000
2008–09 Assessment Score	0.109	0.003	0.000

Table H-14. Grade 5 Math Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	181.233	1.229	0.000
2008–09 Prior Score Missing	72.493	1.776	0.000
Student Poverty Status	–1.307	0.089	0.000
Student SWD Status	–2.639	0.104	0.000
Student ELL Status	0.132	0.159	0.407
2009–10 Assessment Score	0.633	0.003	0.000
2008–09 Assessment Score	0.104	0.003	0.000

Table H-15. Grade 6 Math Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	77.057	1.697	0.000
2007–08 Prior Score Missing	107.260	4.319	0.000
2008–09 Prior Score Missing	51.554	2.869	0.000
2009–10 Assessment Score	0.652	0.005	0.000
2008–09 Assessment Score	0.075	0.004	0.000
2007–08 Assessment Score	0.156	0.006	0.000

Table H-16. Grade 6 Math Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	88.728	1.850	0.000
2007–08 Prior Score Missing	100.874	4.362	0.000
2008–09 Prior Score Missing	50.512	2.866	0.000
Student Poverty status	–1.009	0.084	0.000
Student SWD status	–2.365	0.103	0.000
Student ELL status	2.152	0.167	0.000
2009–10 Assessment Score	0.647	0.005	0.000
2008–09 Assessment Score	0.074	0.004	0.000
2007–08 Assessment Score	0.147	0.006	0.000

Table H-17. Grade 7 Math Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	105.155	1.186	0.000
2007–08 Prior Score Missing	28.926	1.884	0.000
2008–09 Prior Score Missing	46.418	2.844	0.000
2009–10 Assessment Score	0.735	0.004	0.000
2008–09 Assessment Score	0.066	0.004	0.000
2007–08 Assessment Score	0.042	0.003	0.000

Table H-18. Grade 7 Math Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	103.337	1.320	0.000
2007–08 Prior Score Missing	28.733	1.884	0.000
2008–09 Prior Score Missing	46.119	2.843	0.000
Student Poverty status	–0.776	0.068	0.000
Student SWD status	0.256	0.086	0.003
Student ELL status	2.627	0.153	0.000
2009–10 Assessment Score	0.738	0.005	0.000
2008–09 Assessment Score	0.066	0.004	0.000
2007–08 Assessment Score	0.042	0.003	0.000

Table H-19. Grade 8 Math Model Coefficients—Unadjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	102.513	1.618	0.000
2007–08 Prior Score Missing	–8.191	3.010	0.007
2008–09 Prior Score Missing	51.278	4.575	0.000
2009–10 Assessment Score	0.791	0.007	0.000
2008–09 Assessment Score	0.075	0.007	0.000
2007–08 Assessment Score	–0.015	0.004	0.001

Table H-20. Grade 8 Math Model Coefficients—Adjusted Model

Effect Name	Effect	Standard Error	Significance
Constant Term	96.867	1.763	0.000
2007–08 Prior Score Missing	–8.198	3.000	0.006
2008–09 Prior Score Missing	53.017	4.563	0.000
Student Poverty status	–0.078	0.090	0.389
Student SWD status	0.113	0.117	0.333
Student ELL status	5.664	0.207	0.000
2009–10 Assessment Score	0.795	0.007	0.000
2008–09 Assessment Score	0.078	0.007	0.000
2007–08 Assessment Score	–0.014	0.004	0.002

Appendix J. Additional Impact Data (Teacher Level by Subject and Grade, and School Level)

Table J-1. Teacher MGP by Students' Mean Prior Scale Score—ELA Grade 4

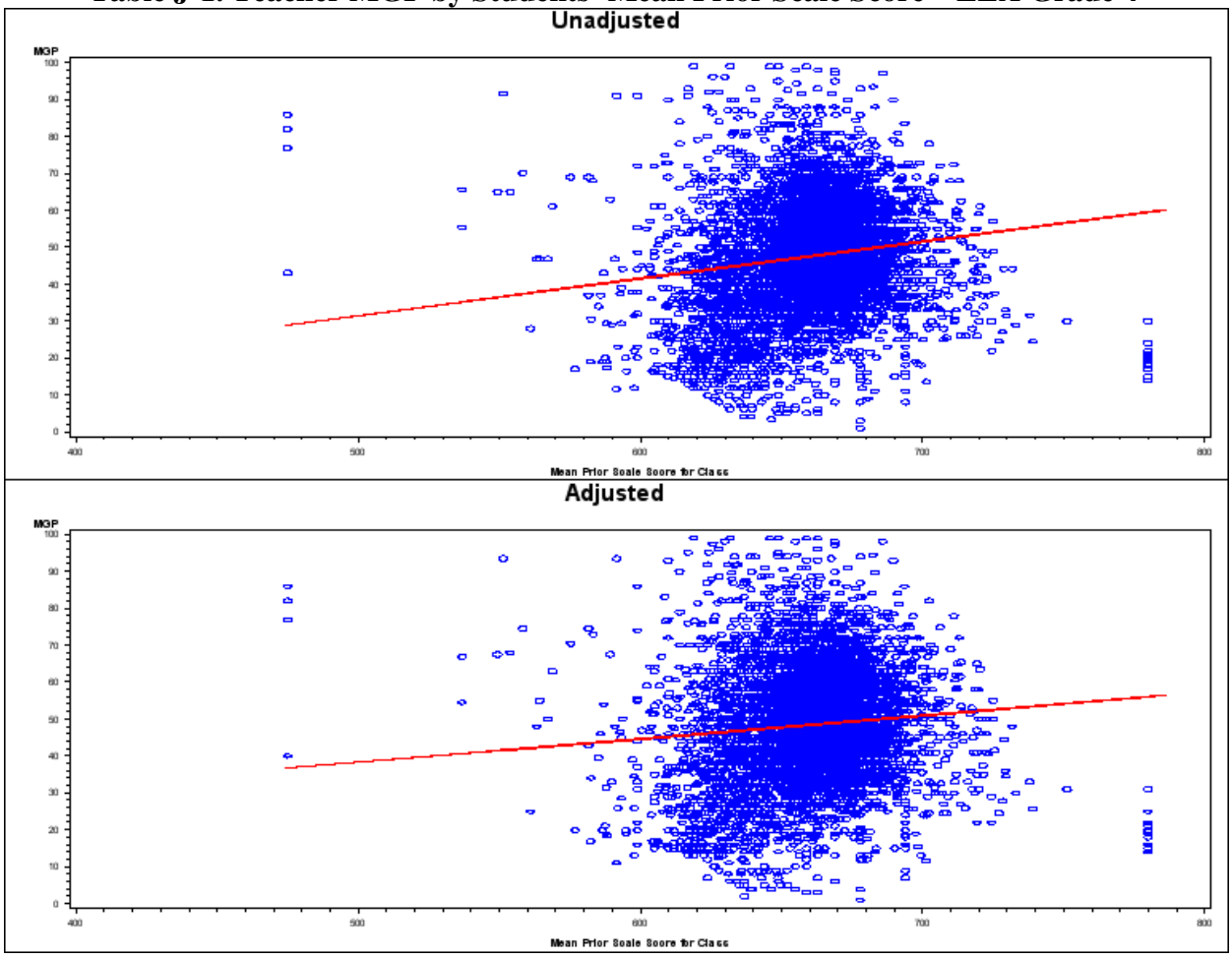


Table J-2. Teacher MGP by Percent SWD—ELA Grade 4

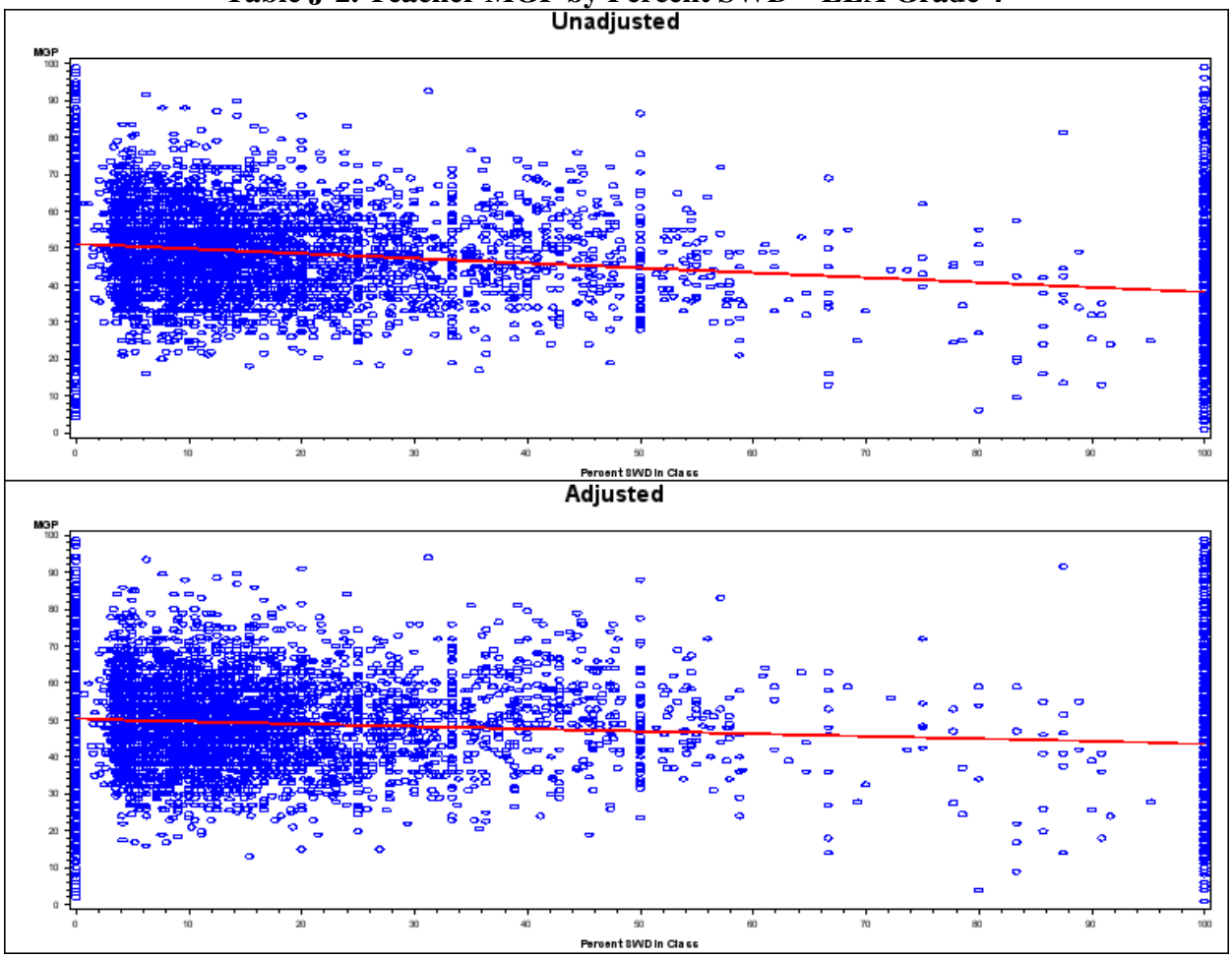


Table J-3. Teacher MGP by Percent ELL—ELA Grade 4

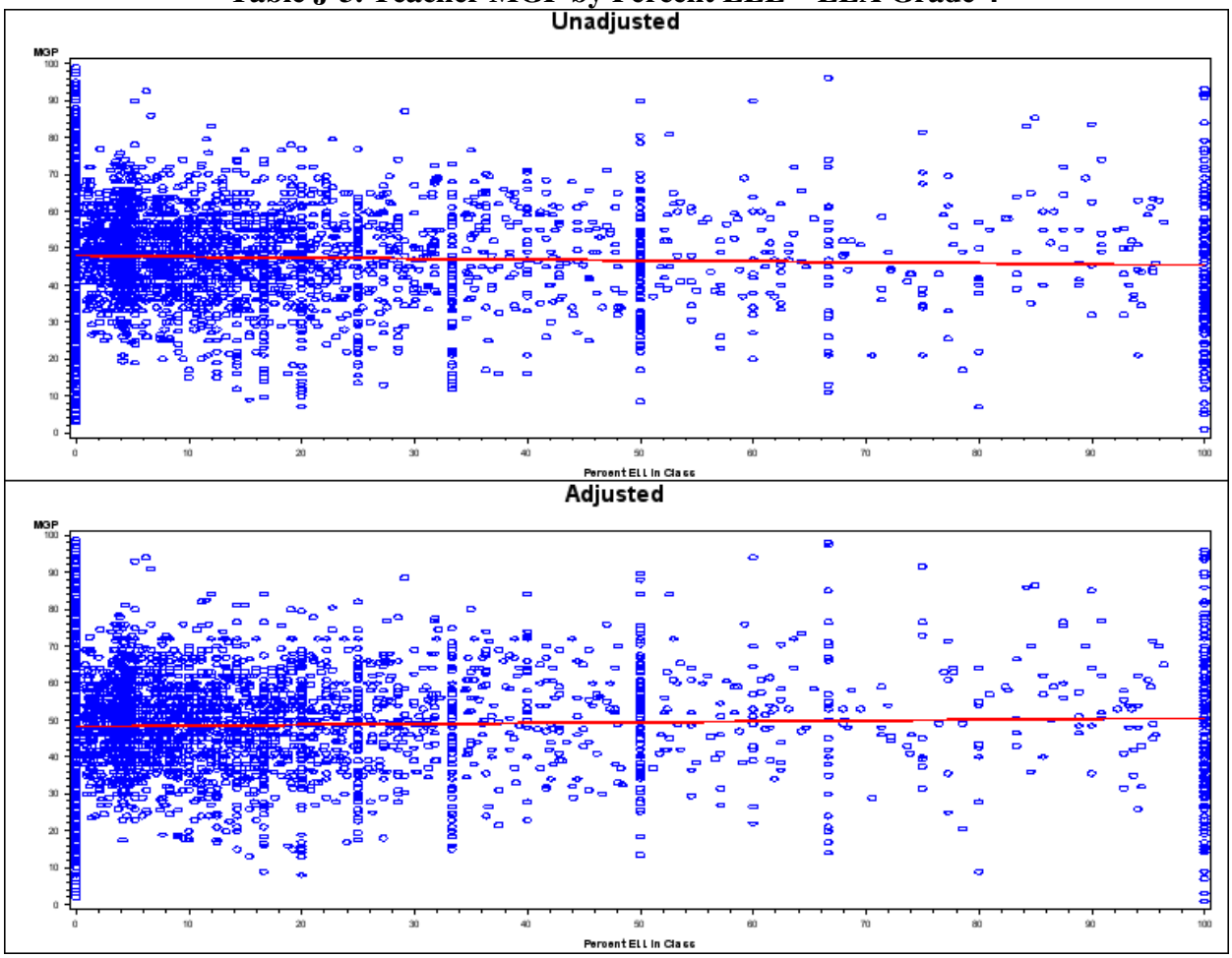


Table J-4. Teacher MGP by Percent Economically Disadvantaged—ELA Grade 4

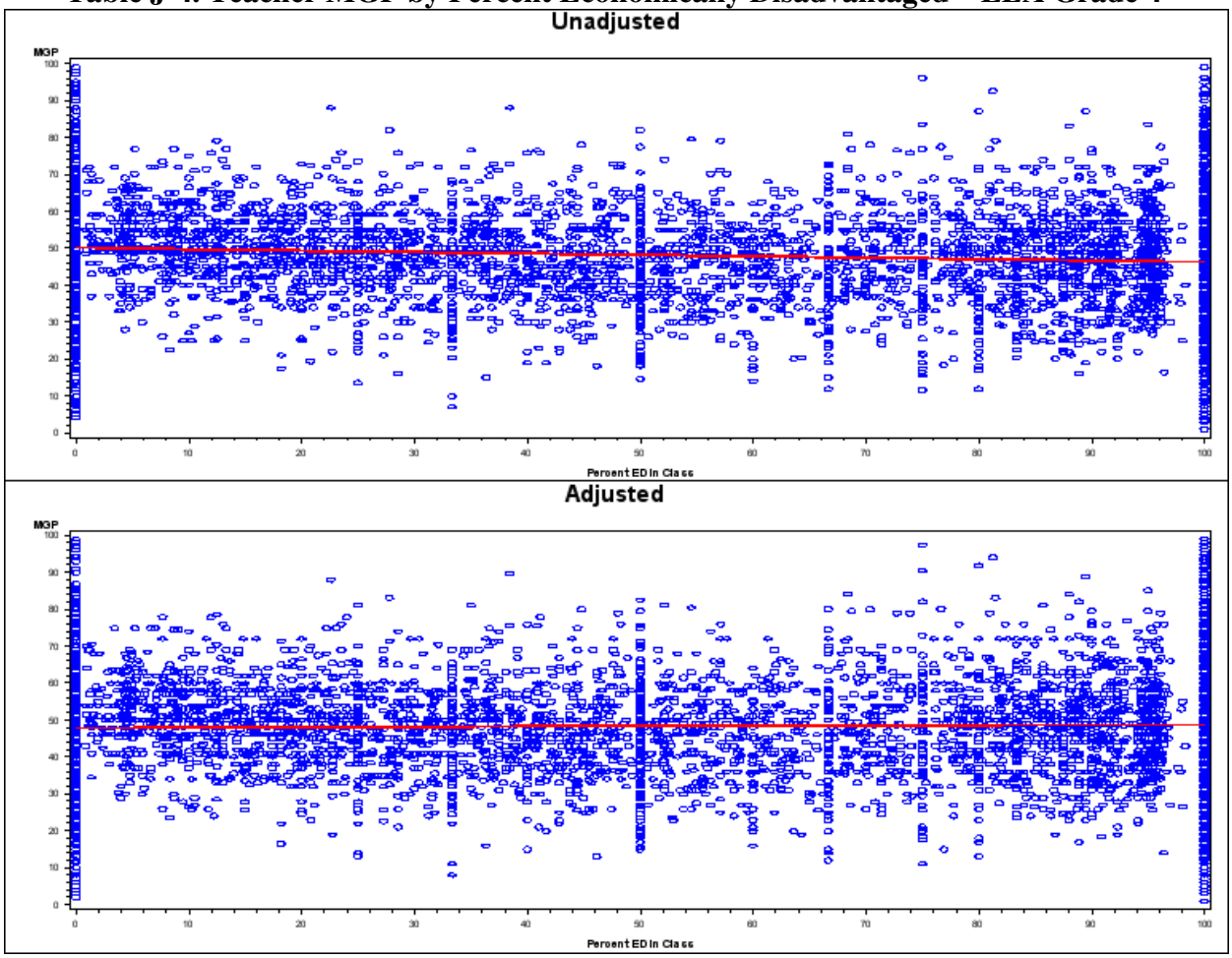


Table J-5. Teacher MGP by Students' Mean Prior Scale Score—ELA Grade 5

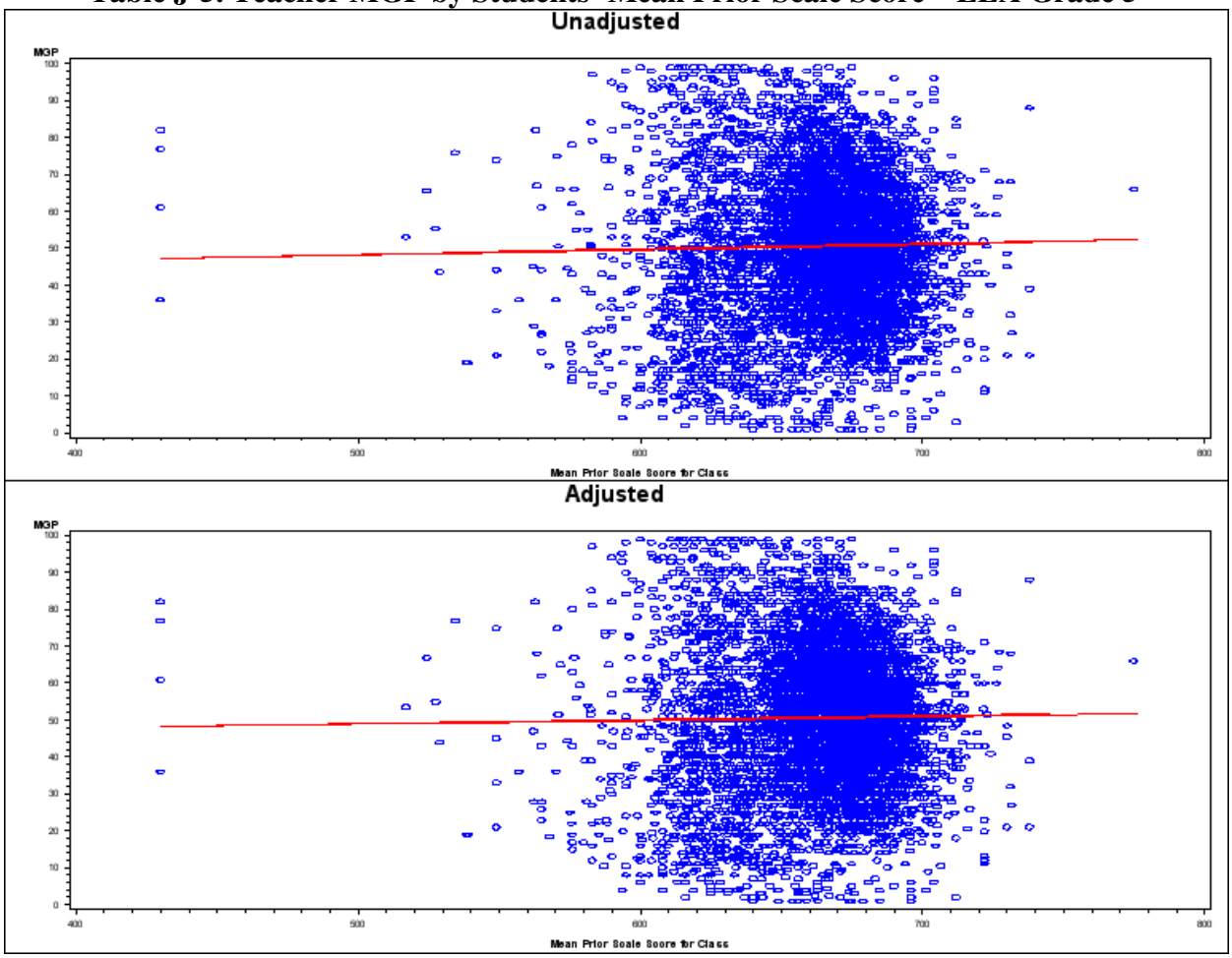


Table J-6. Teacher MGP by Percent SWD—ELA Grade 5

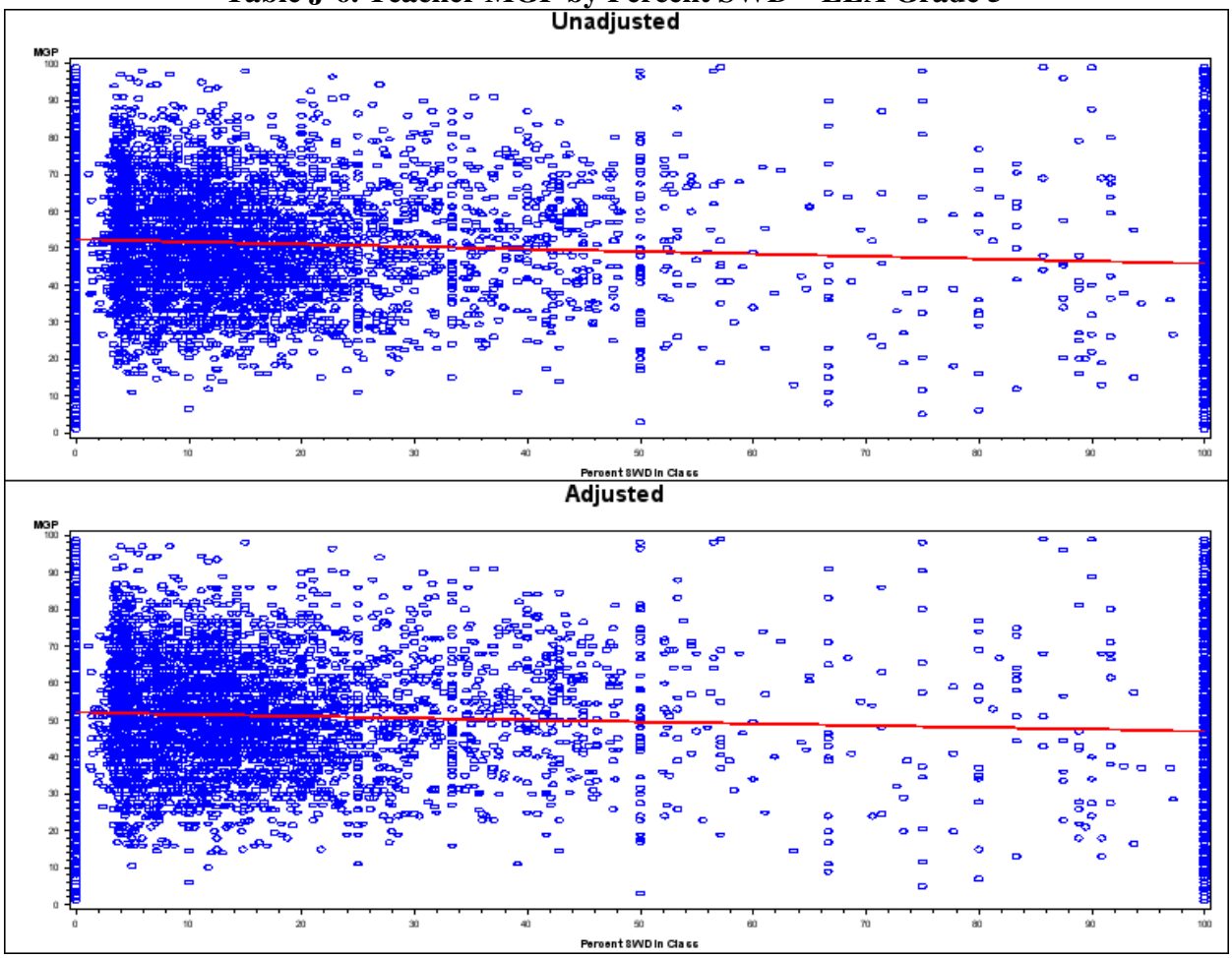


Table J-7. Teacher MGP by Percent ELL—ELA Grade 5

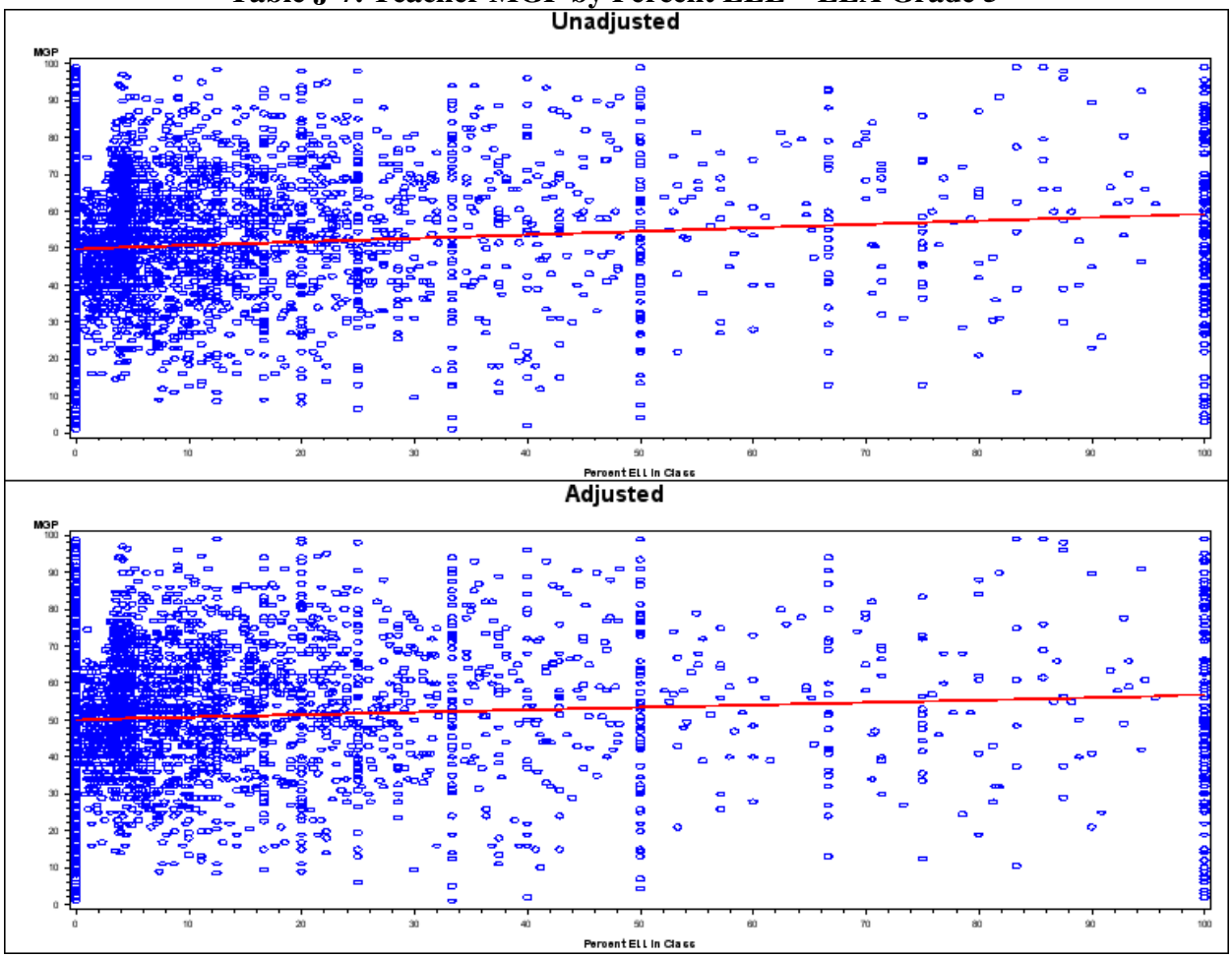


Table J-8. Teacher MGP by Percent Economically Disadvantaged—ELA Grade 5

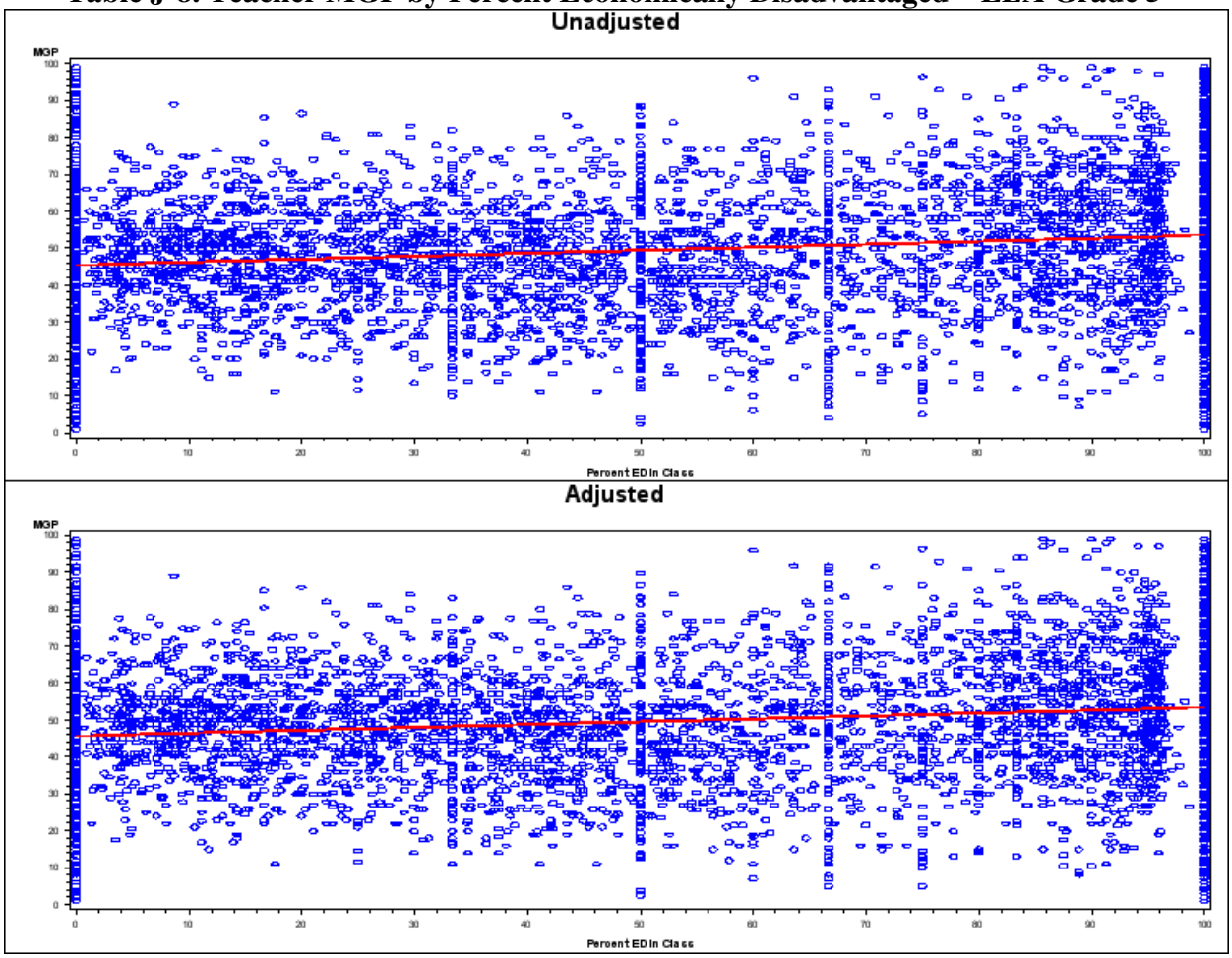


Table J-9. Teacher MGP by Students' Mean Prior Scale Score—ELA Grade 6

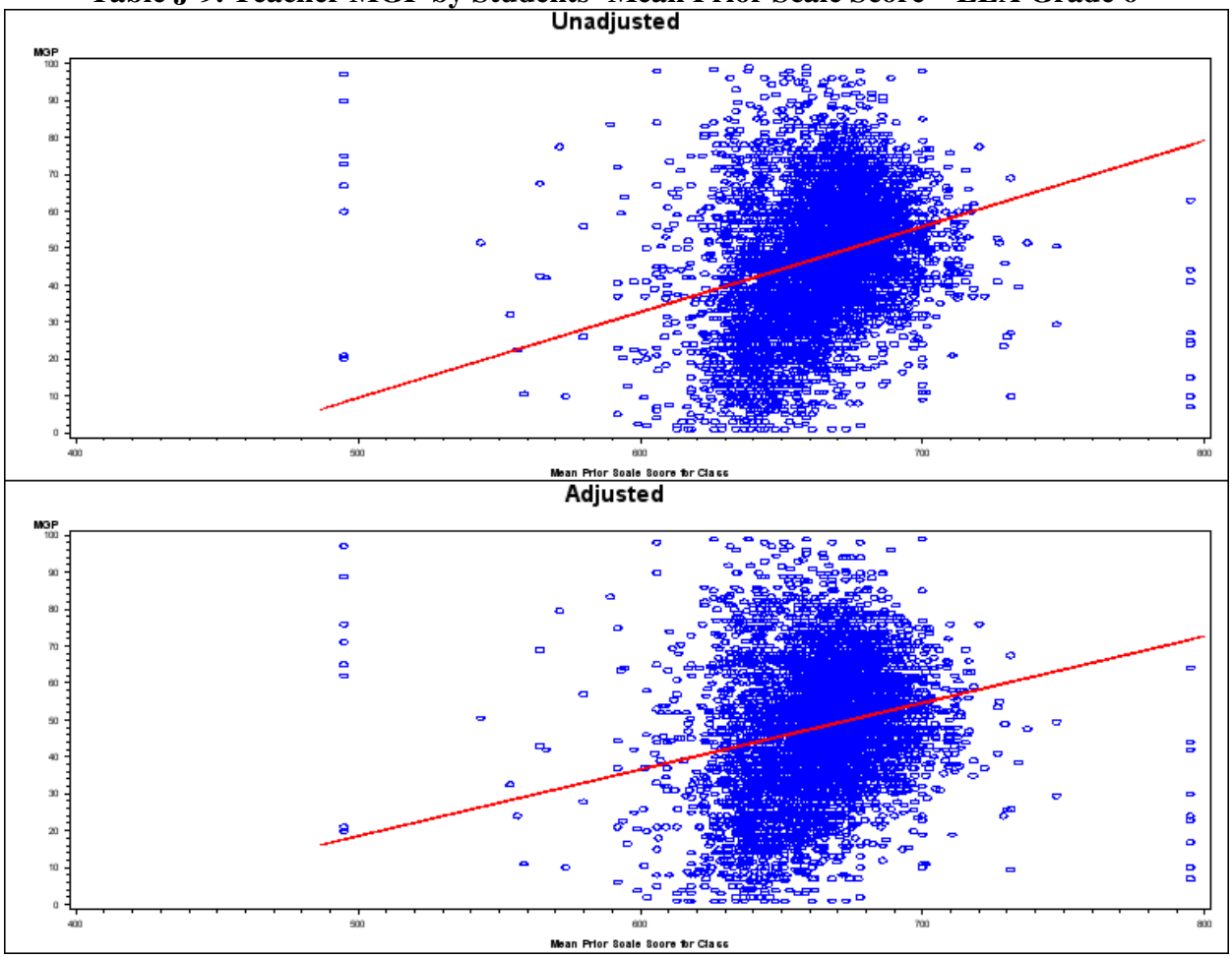


Table J-10. Teacher MGP by Percent SWD—ELA Grade 6

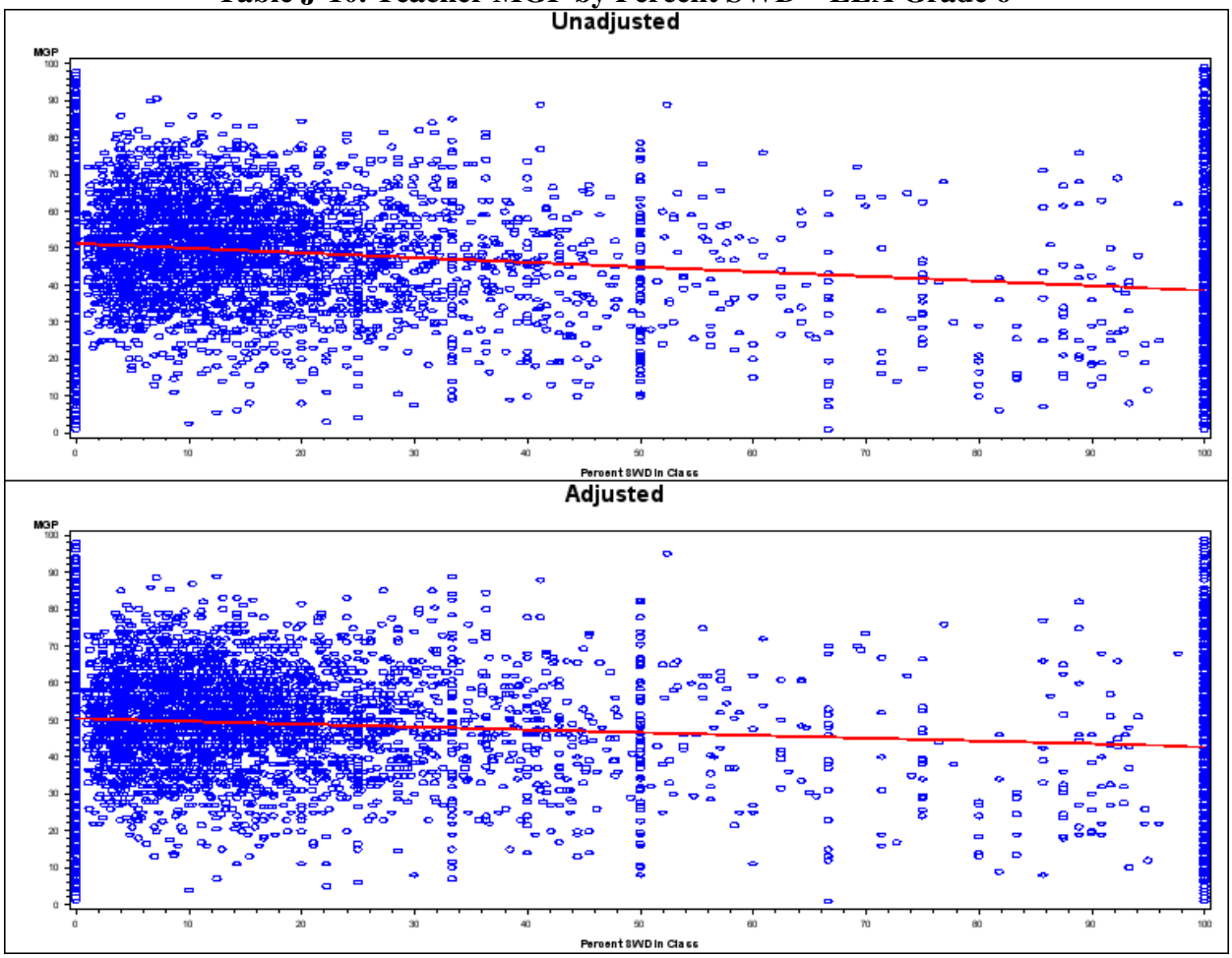


Table J-11. Teacher MGP by Percent ELL—ELA Grade 6

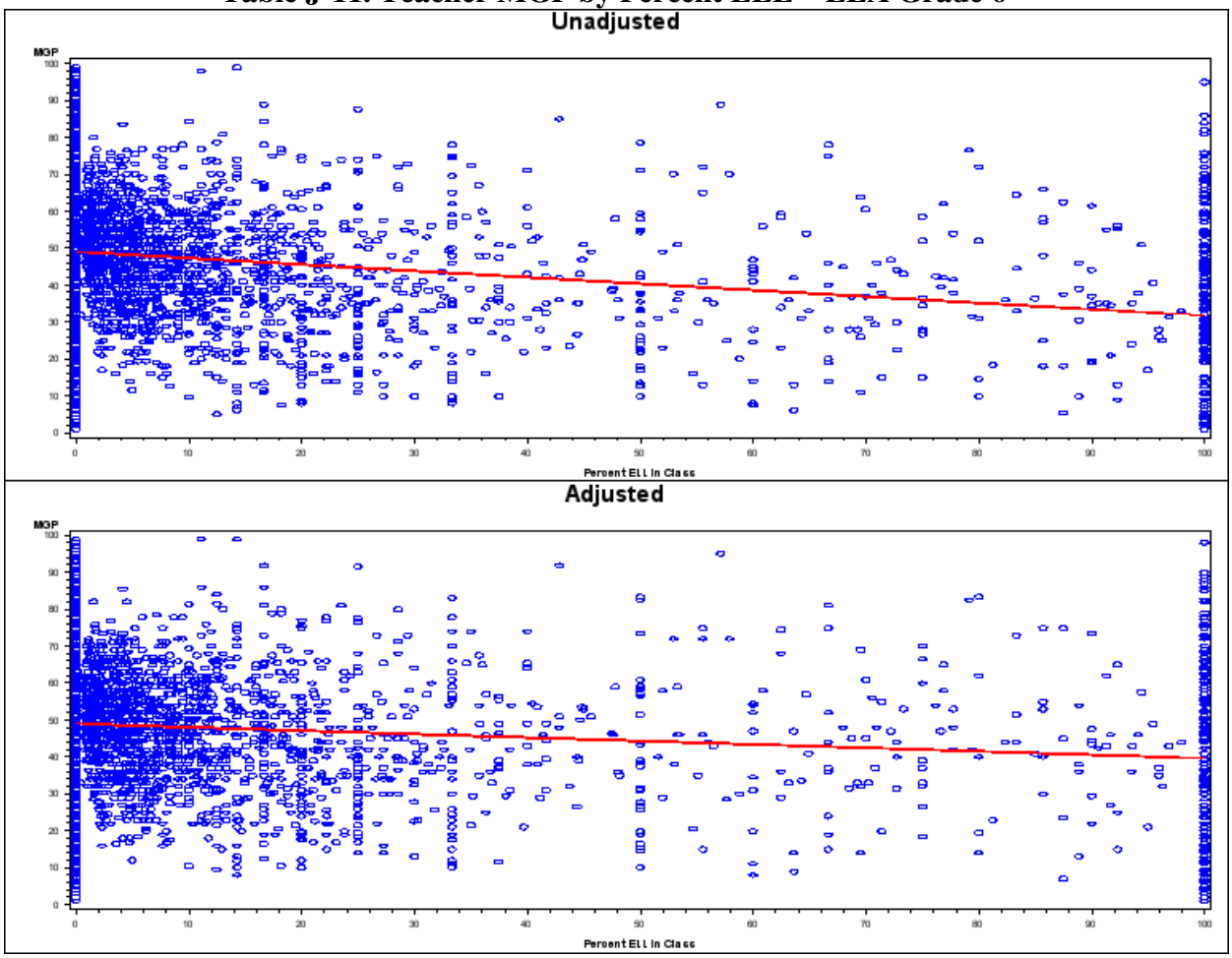


Table J-12. Teacher MGP by Percent Economically Disadvantaged—ELA Grade 6

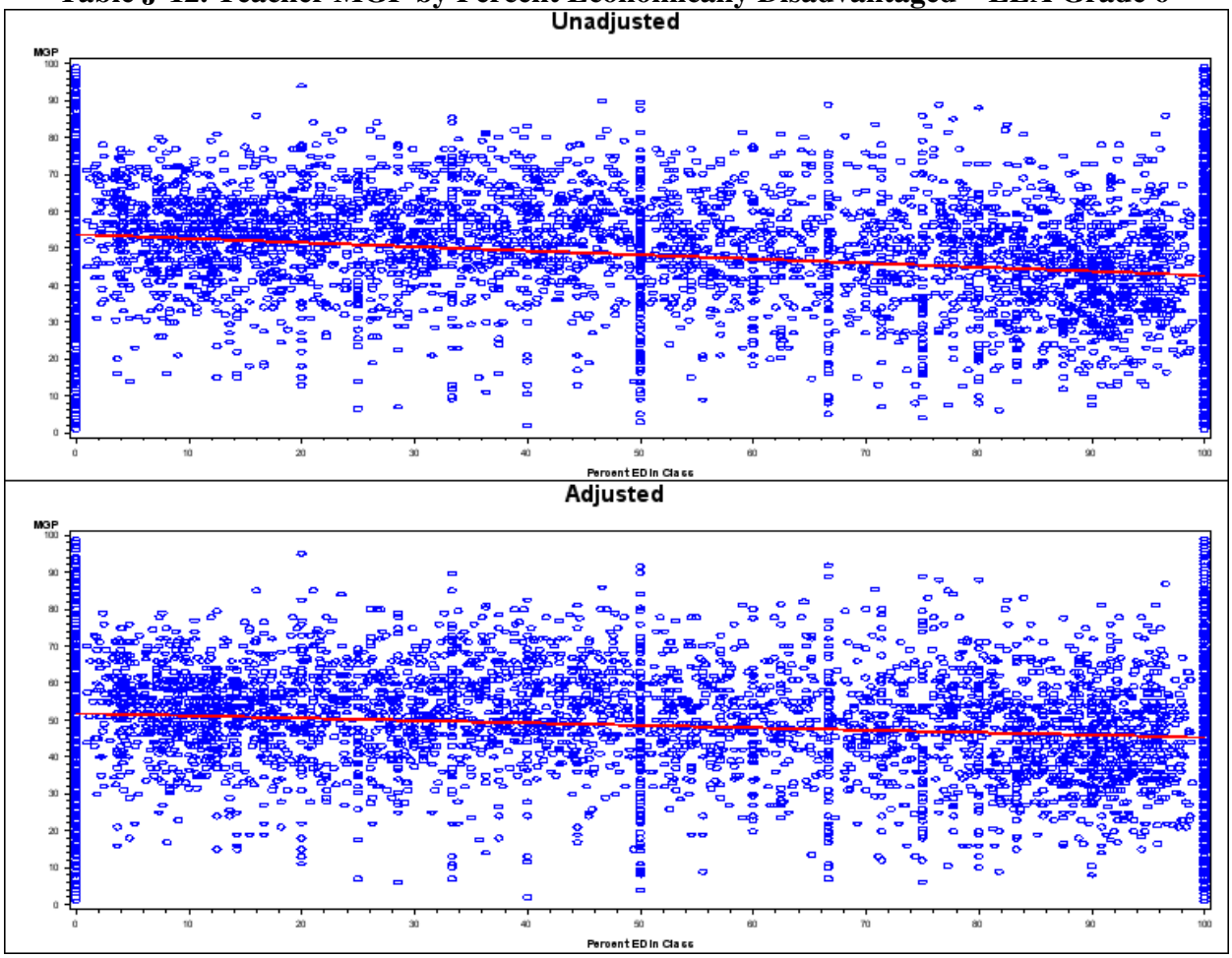


Table J-13. Teacher MGP by Students' Mean Prior Scale Score—ELA Grade 7

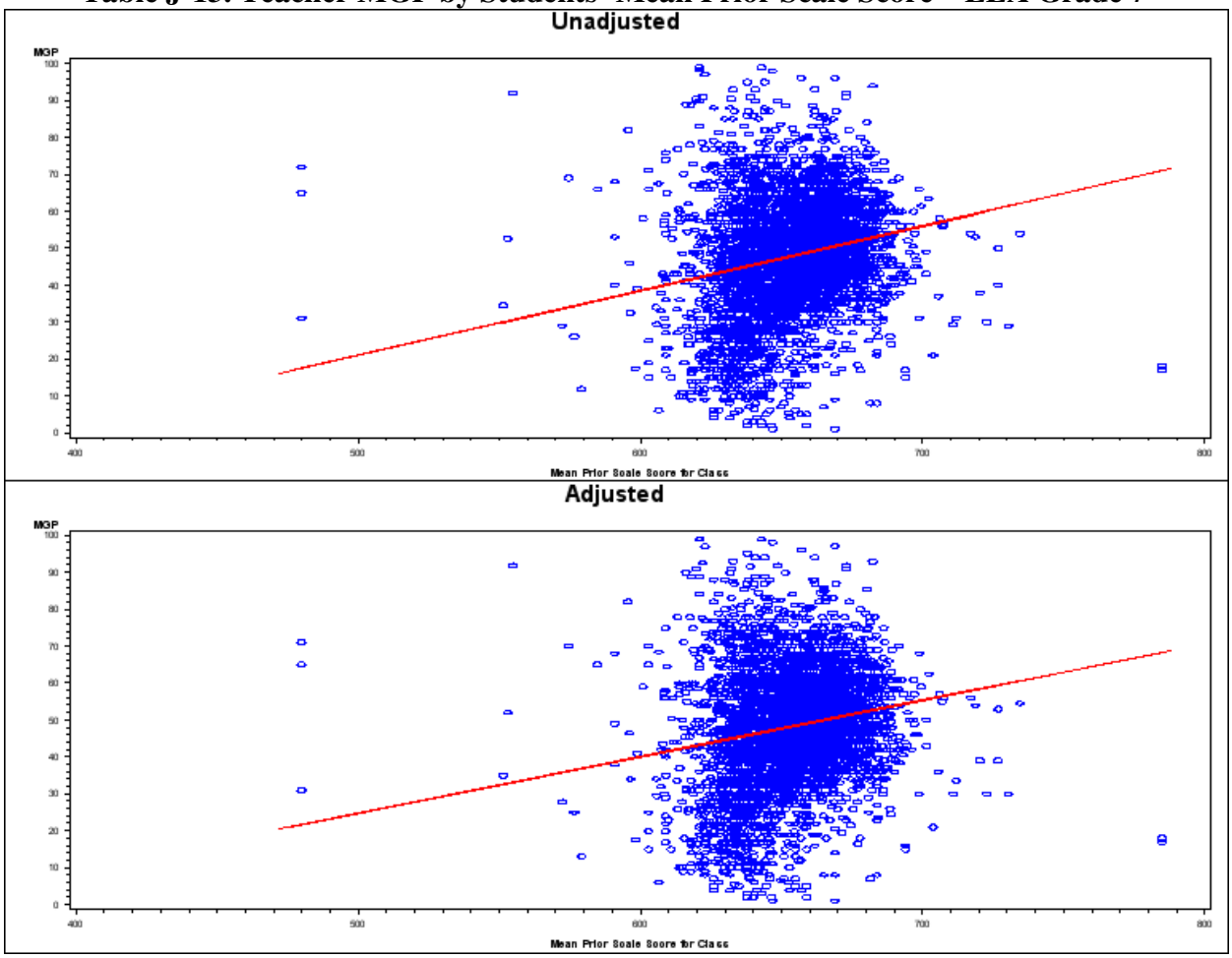


Table J-14. Teacher MGP by Percent SWD—ELA Grade 7

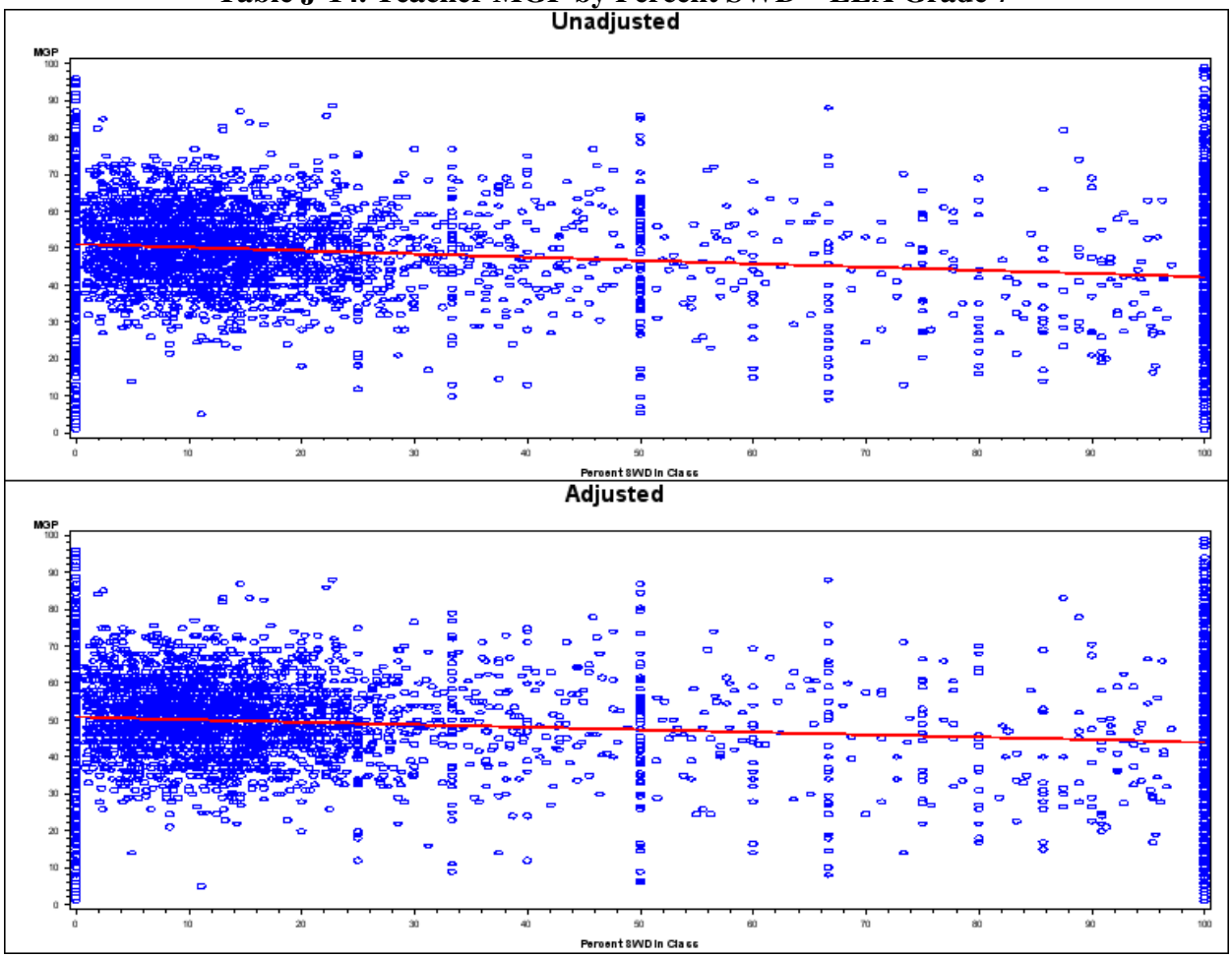


Table J-15. Teacher MGP by Percent ELL—ELA Grade 7

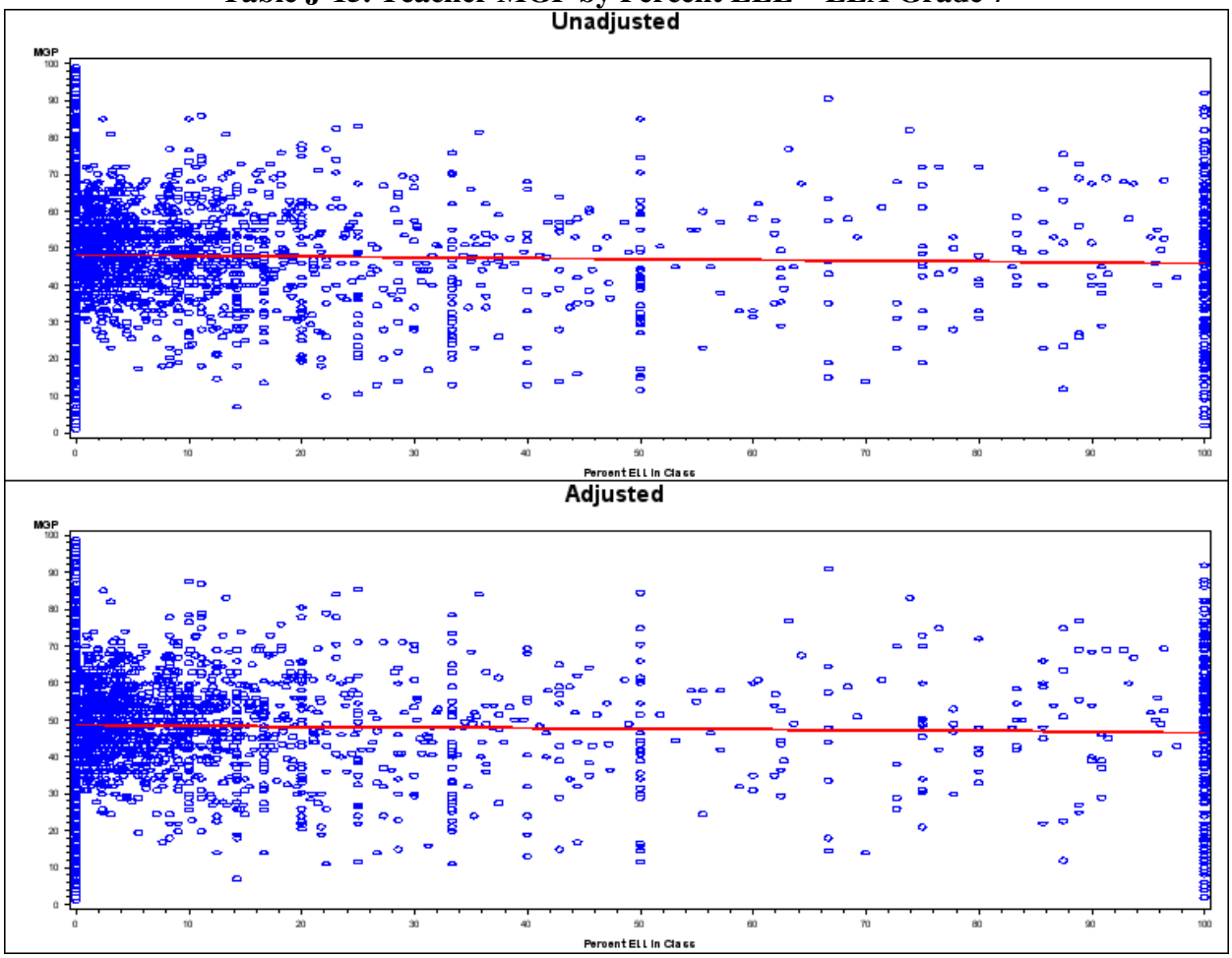


Table J-16. Teacher MGP by Percent Economically Disadvantaged—ELA Grade 7

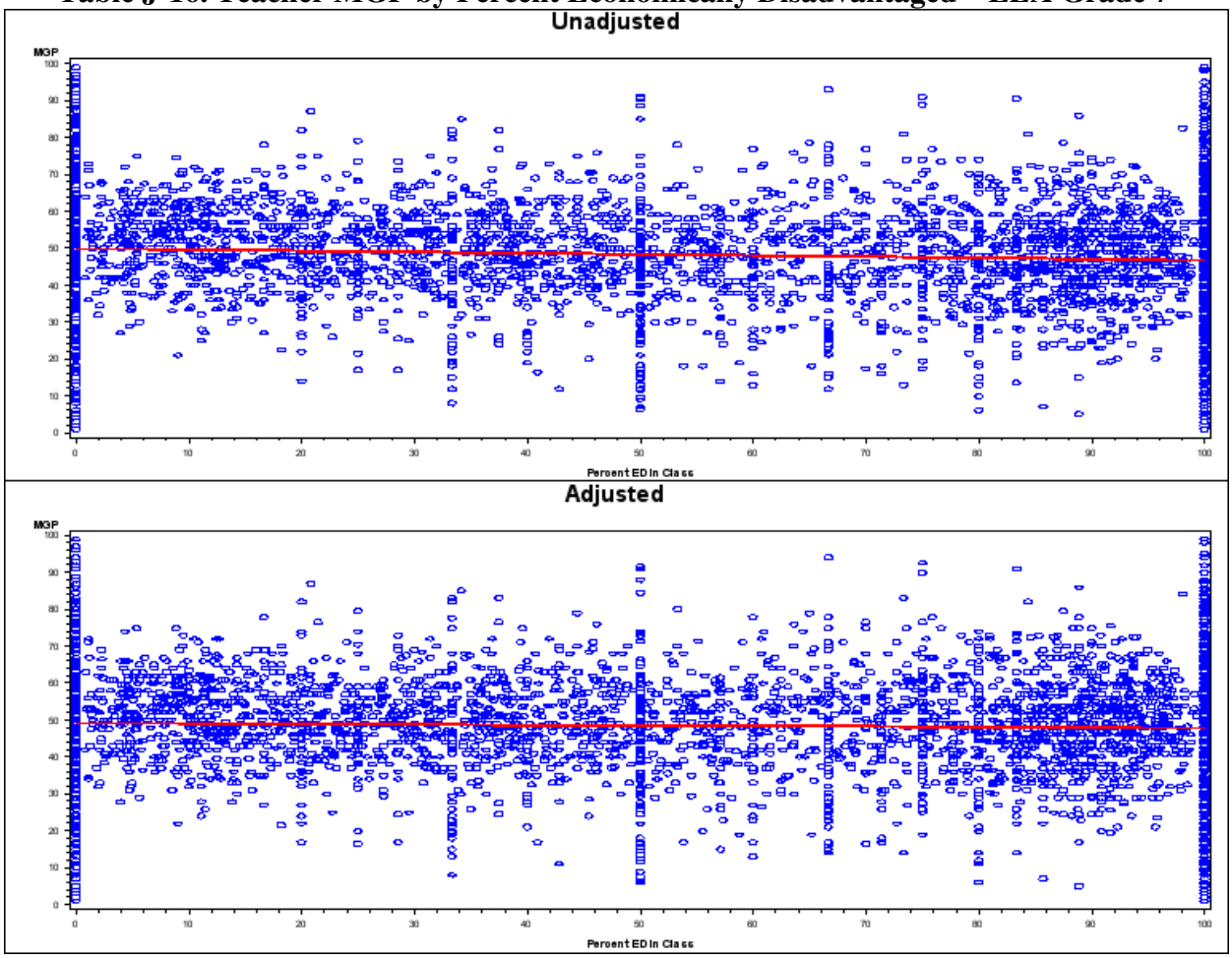


Table J-17. Teacher MGP by Students' Mean Prior Scale Score—ELA Grade 8

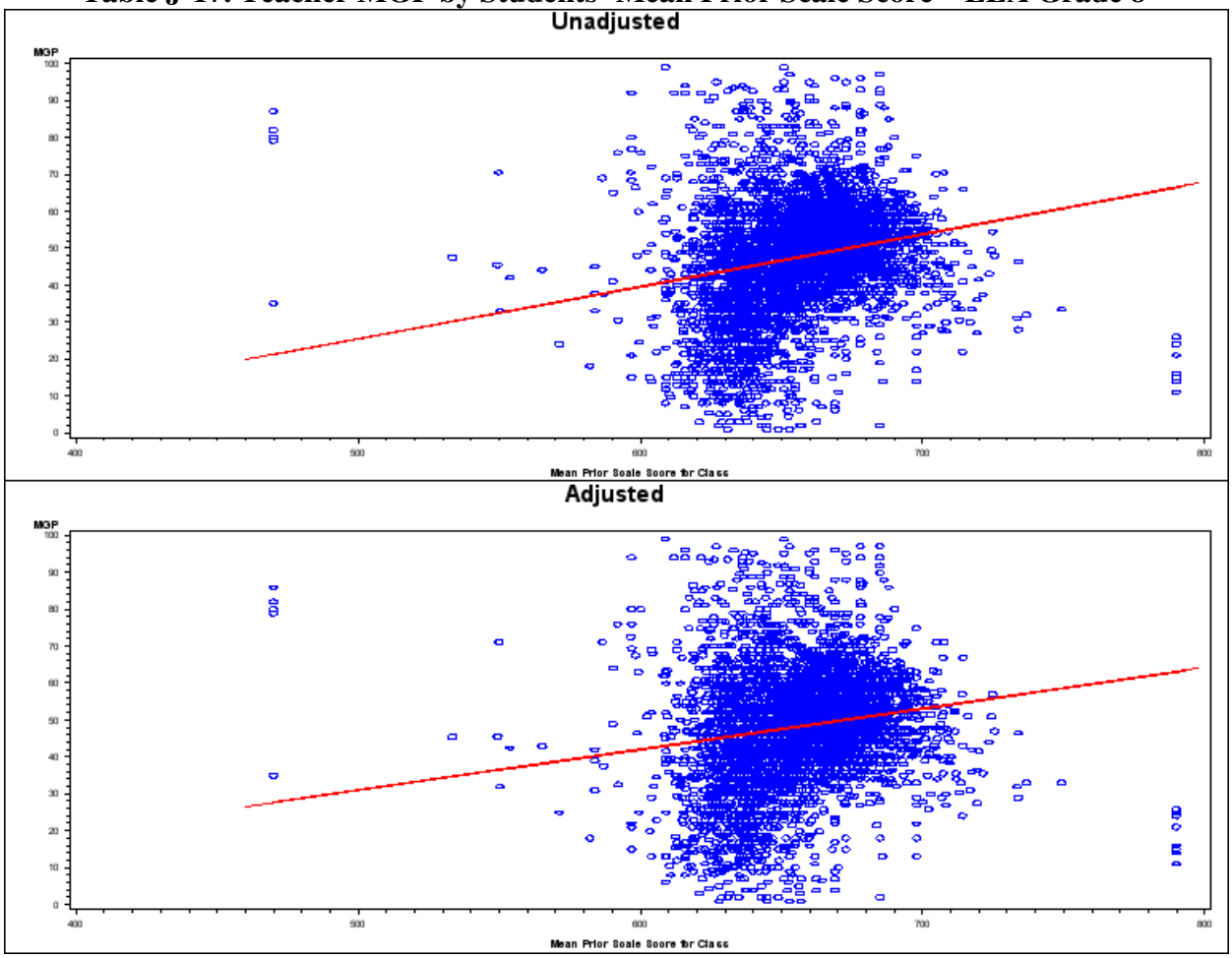


Table J-18. Teacher MGP by Percent SWD—ELA Grade 8

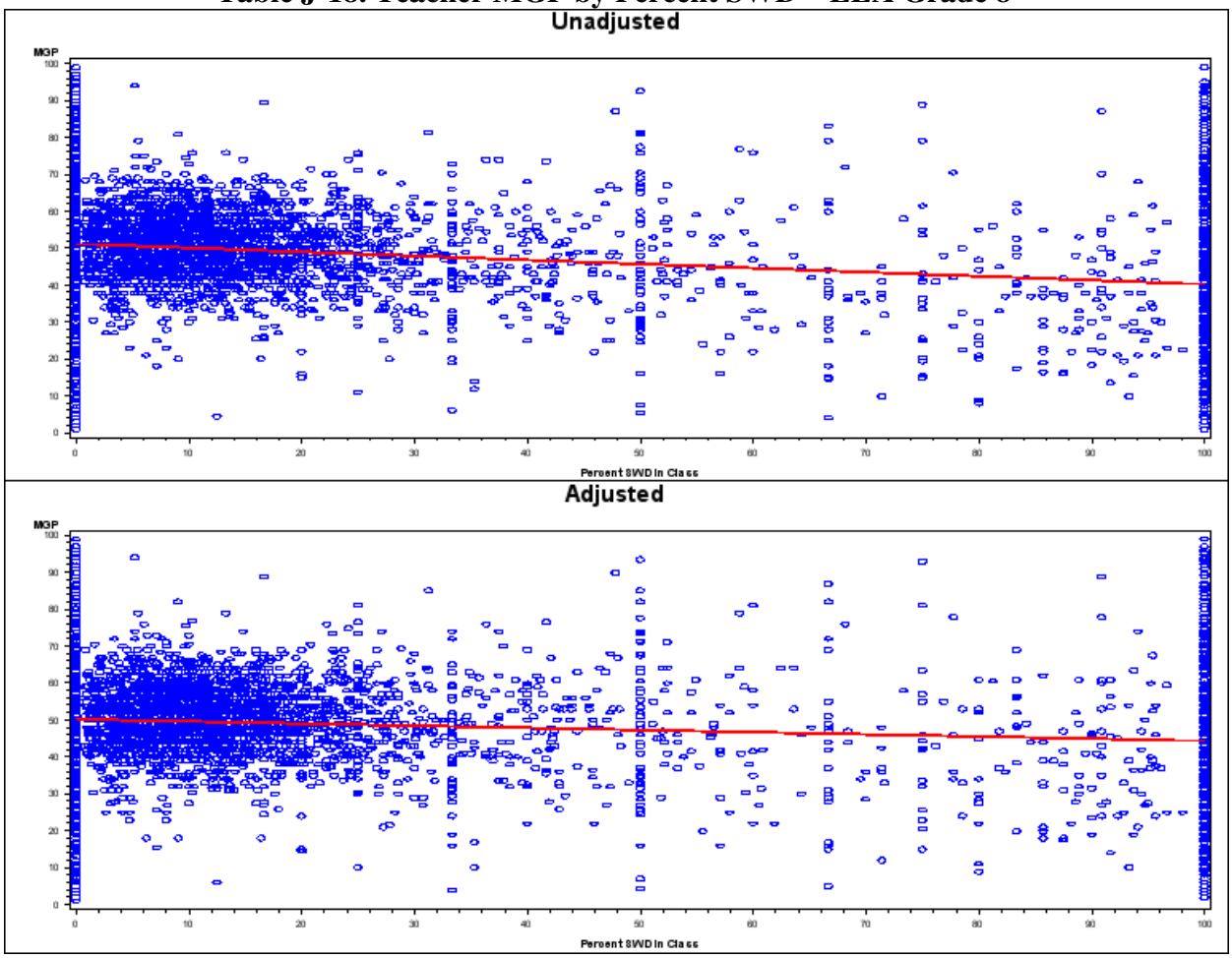


Table J-19. Teacher MGP by Percent ELL—ELA Grade 8

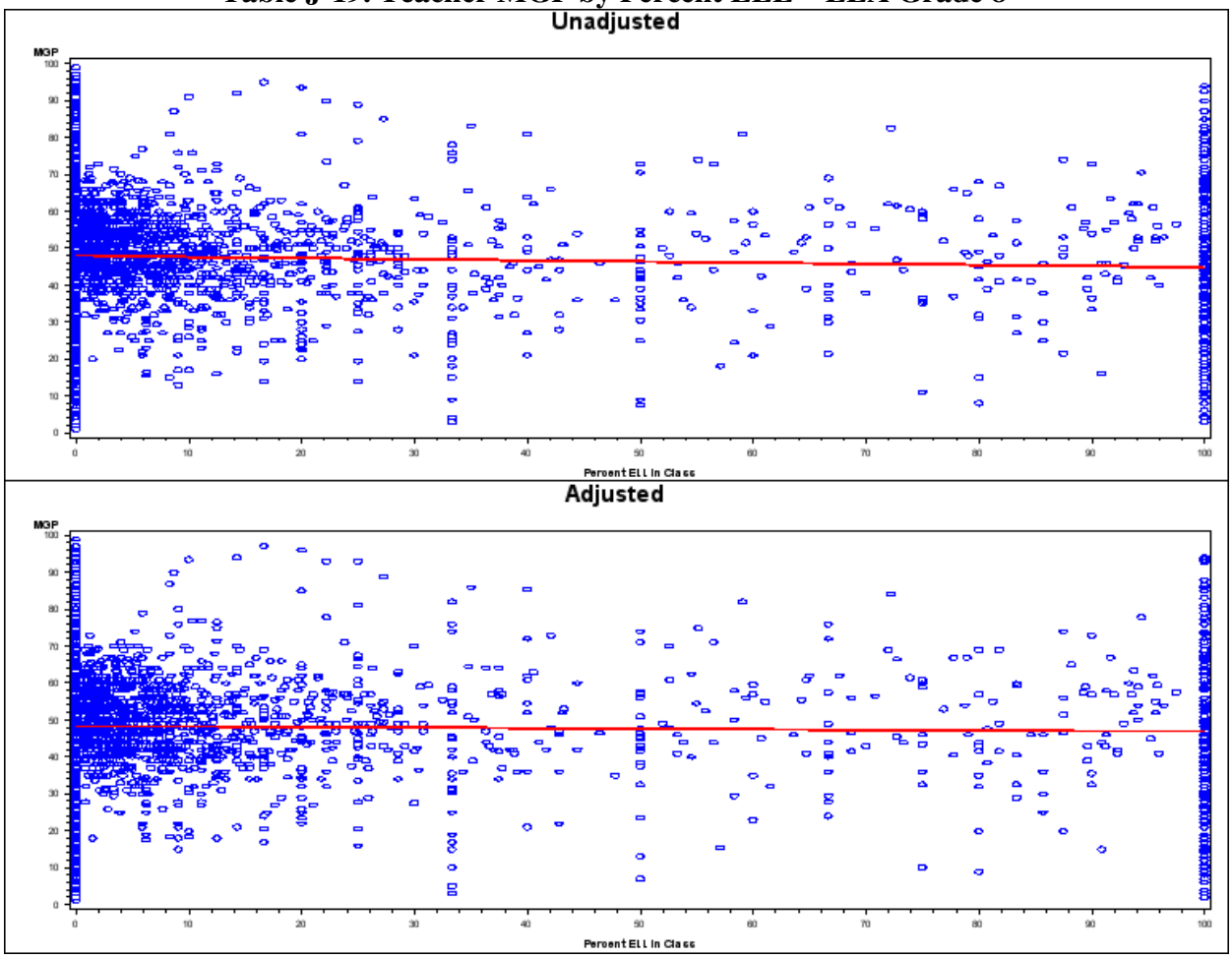


Table J-20. Teacher MGP by Percent Economically Disadvantaged—ELA Grade 8

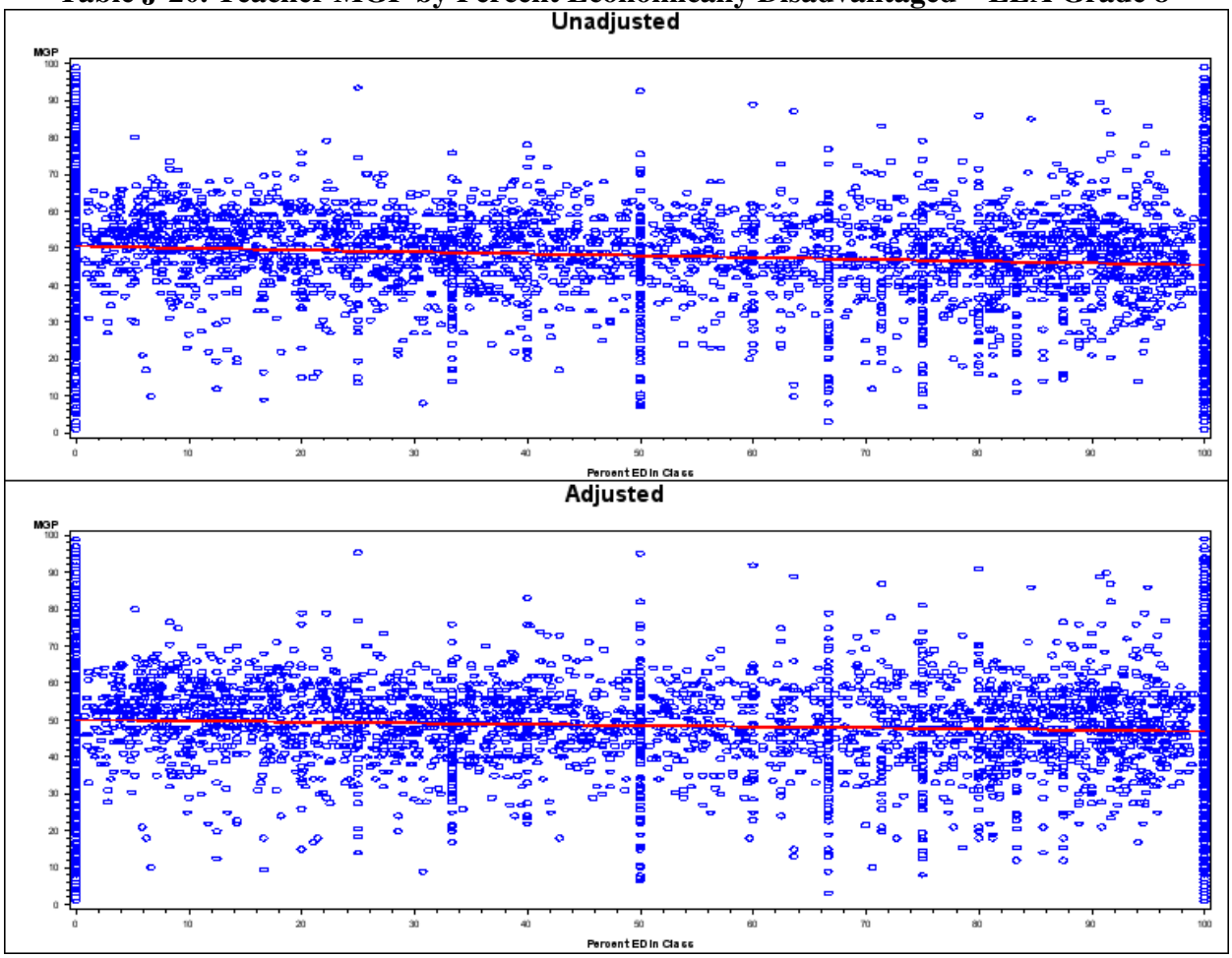


Table J-21. Teacher MGP by Students' Mean Prior Scale Score—Math Grade 4

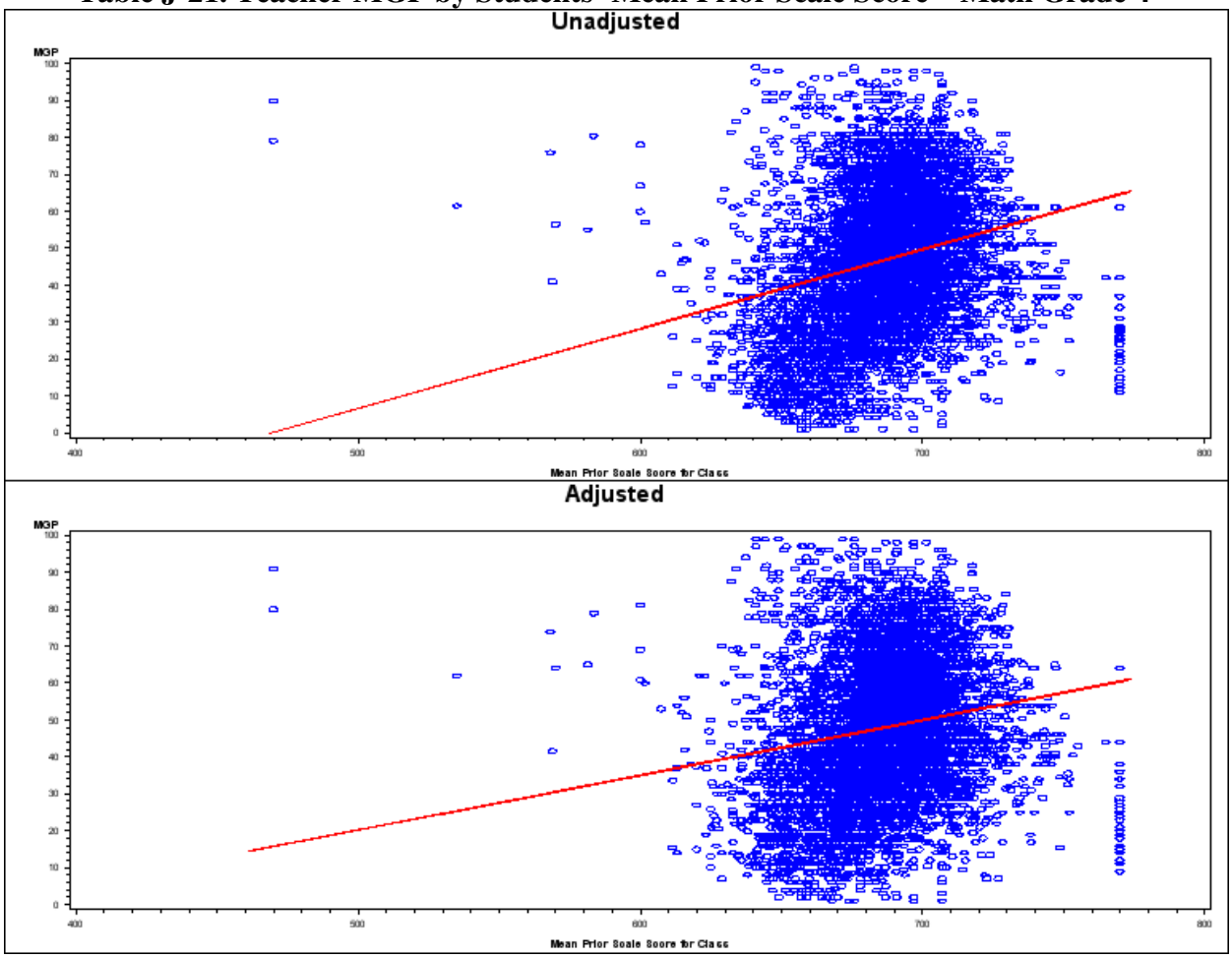


Table J-22. Teacher MGP by Percent SWD—Math Grade 4

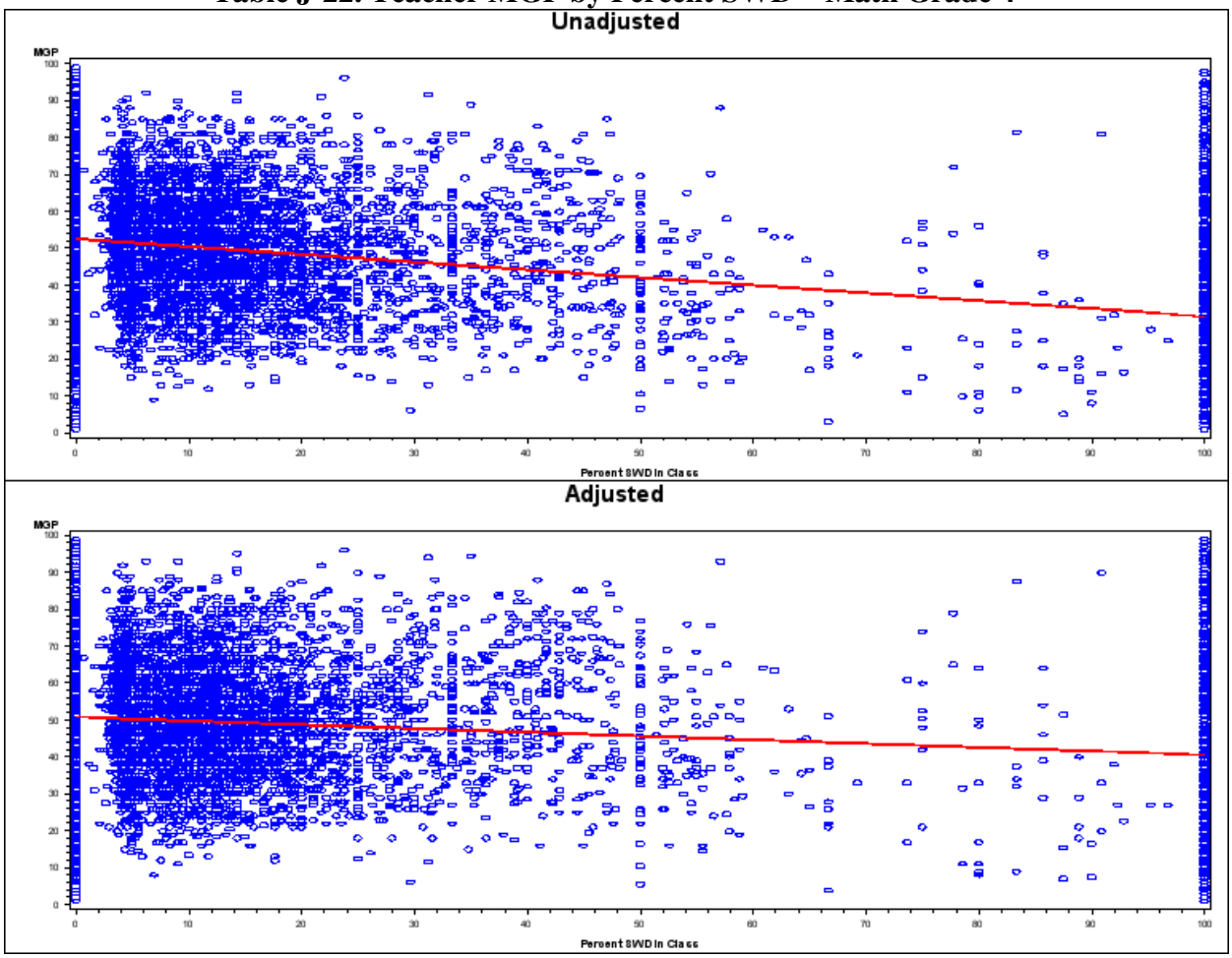


Table J-23. Teacher MGP by Percent ELL—Math Grade 4

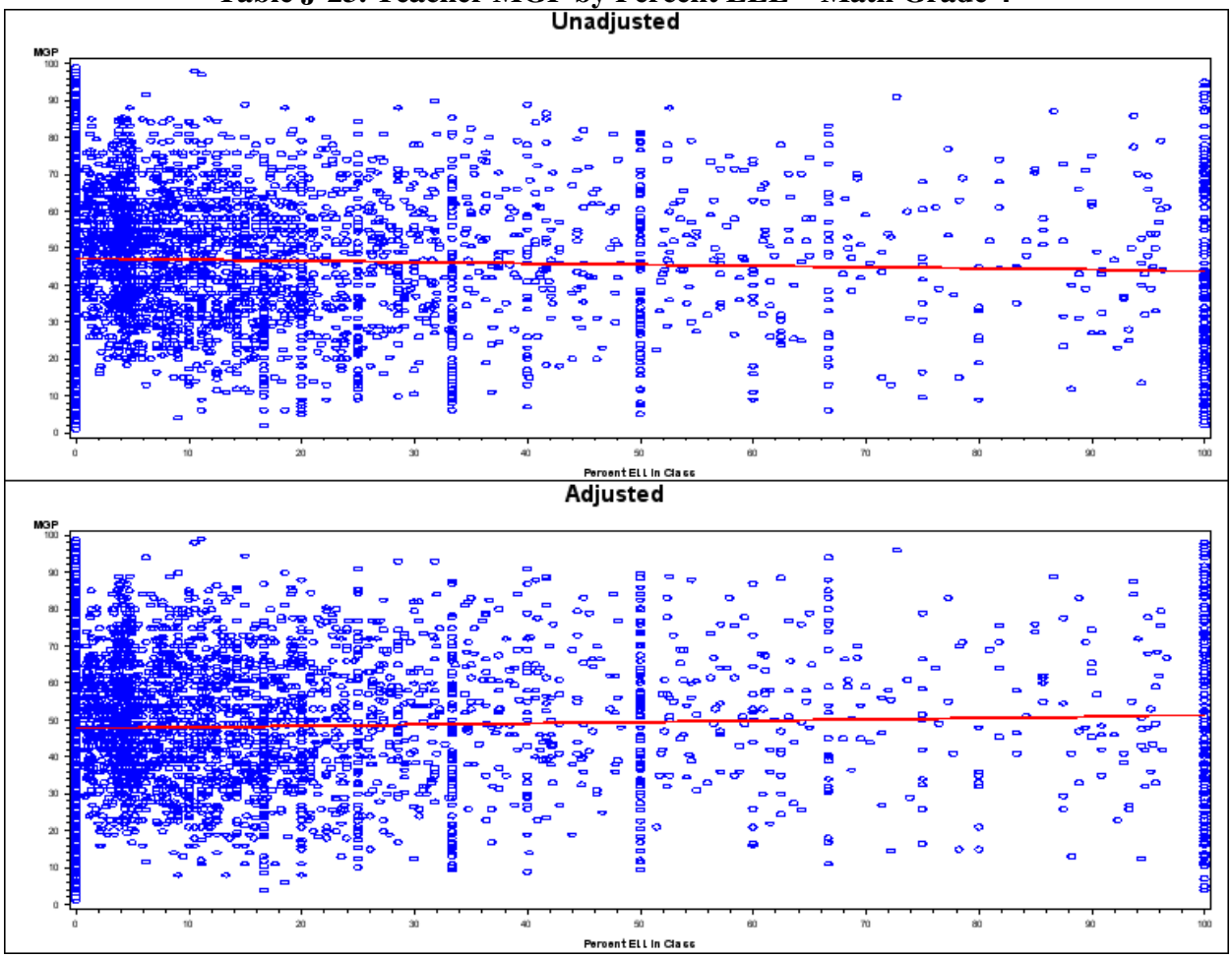


Table J-24. Teacher MGP by Percent Economically Disadvantaged—Math Grade 4

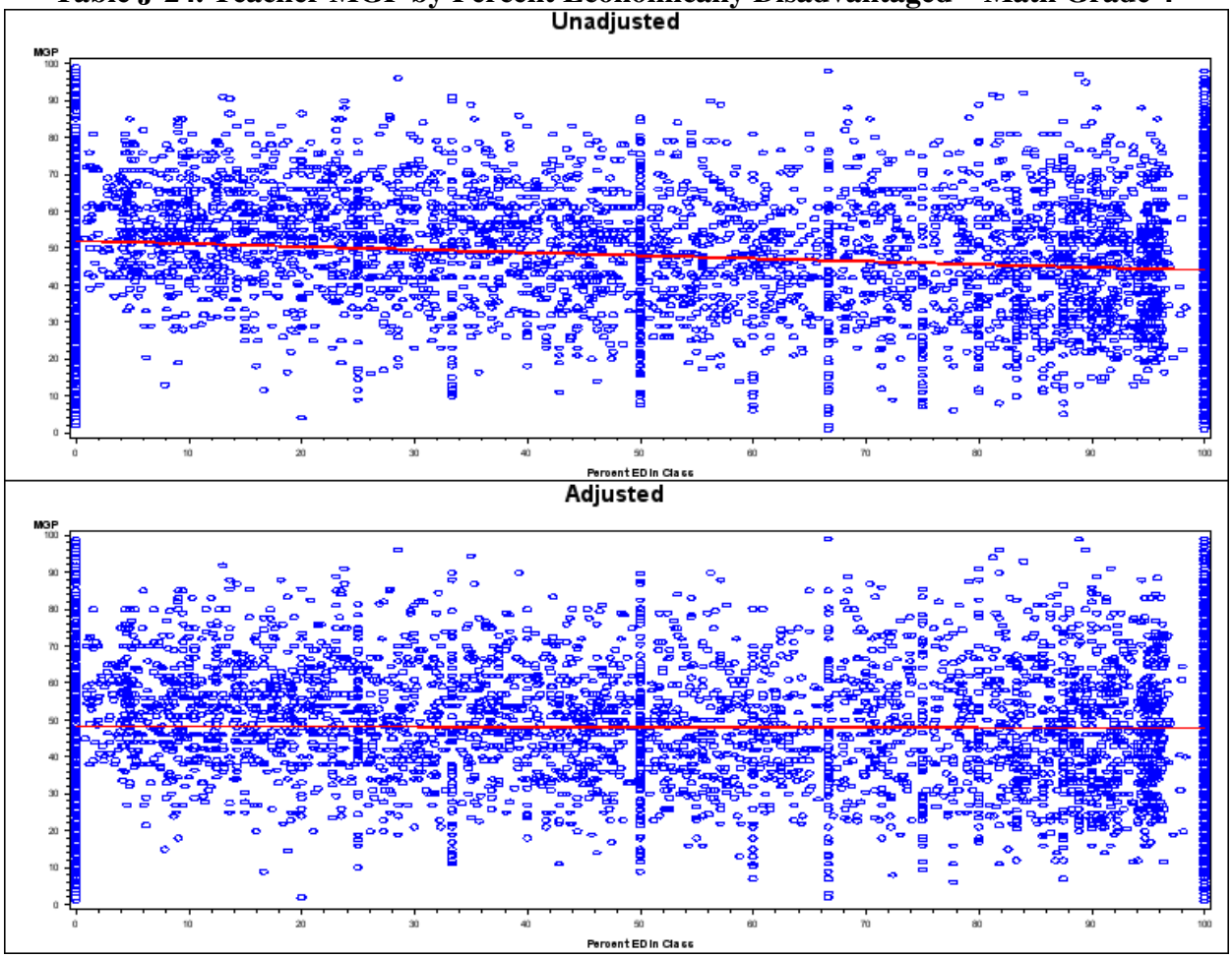


Table J-25. Teacher MGP by Students' Mean Prior Scale Score—Math Grade 5

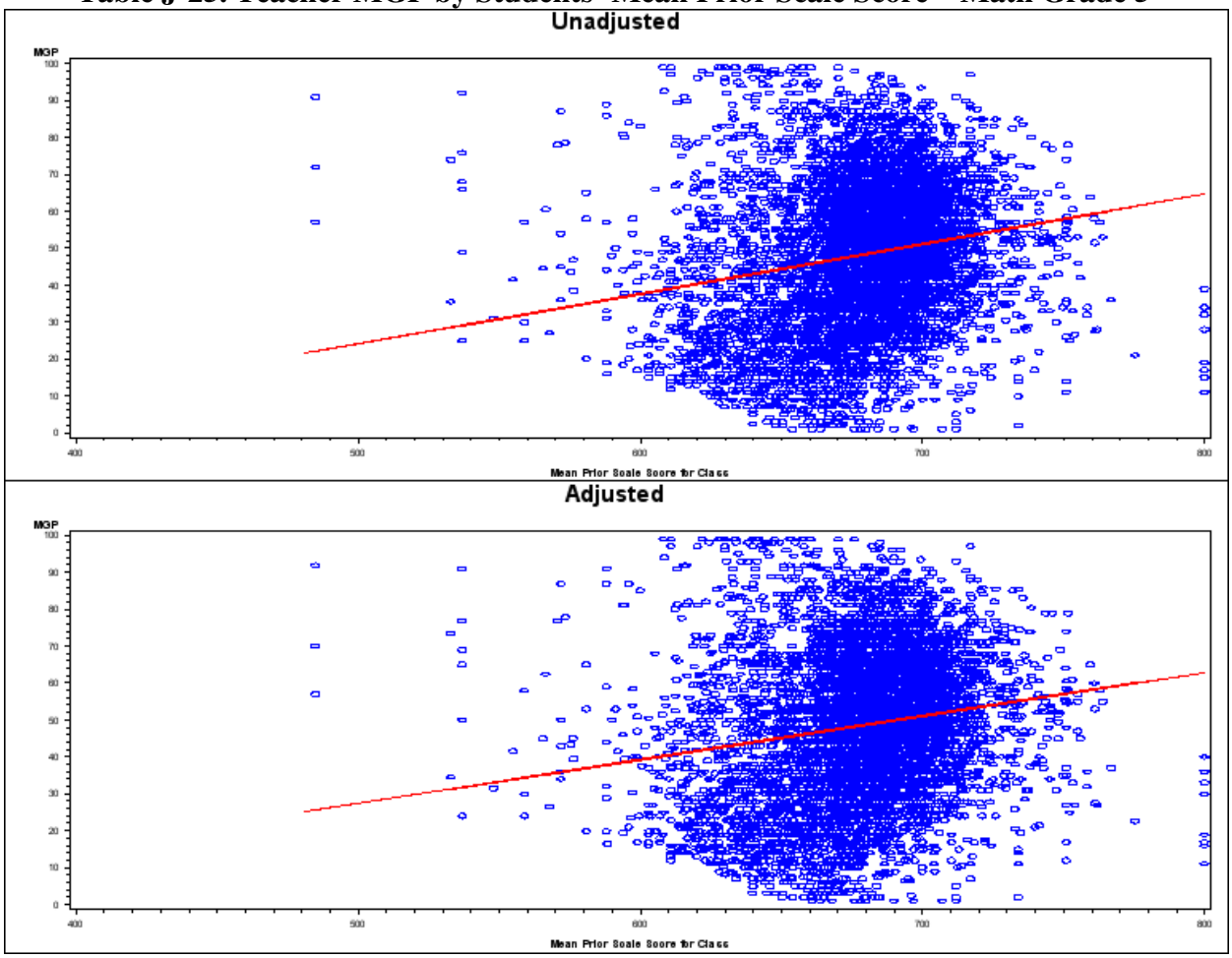


Table J-26. Teacher MGP by Percent SWD—Math Grade 5

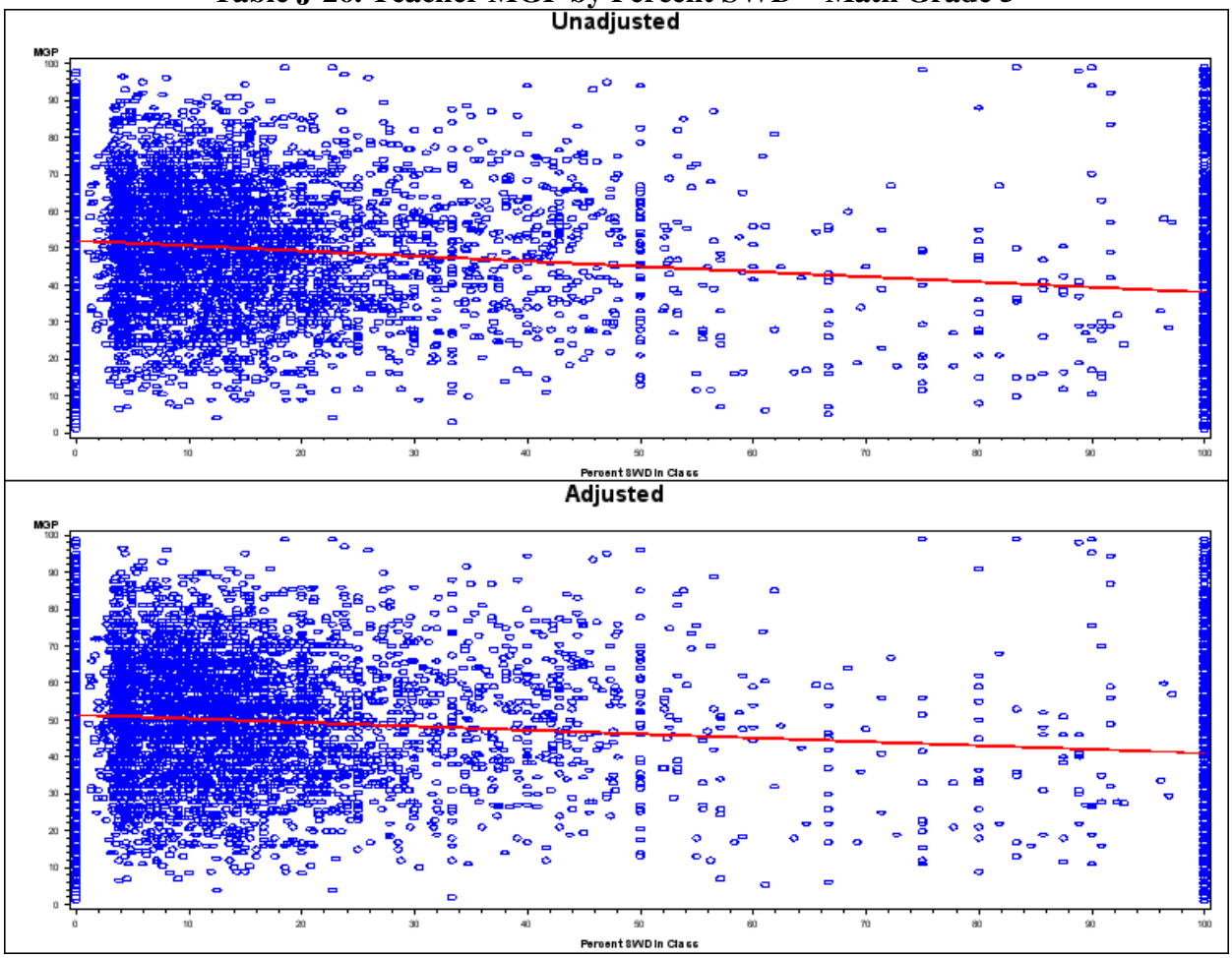


Table J-27. Teacher MGP by Percent ELL—Math Grade 5

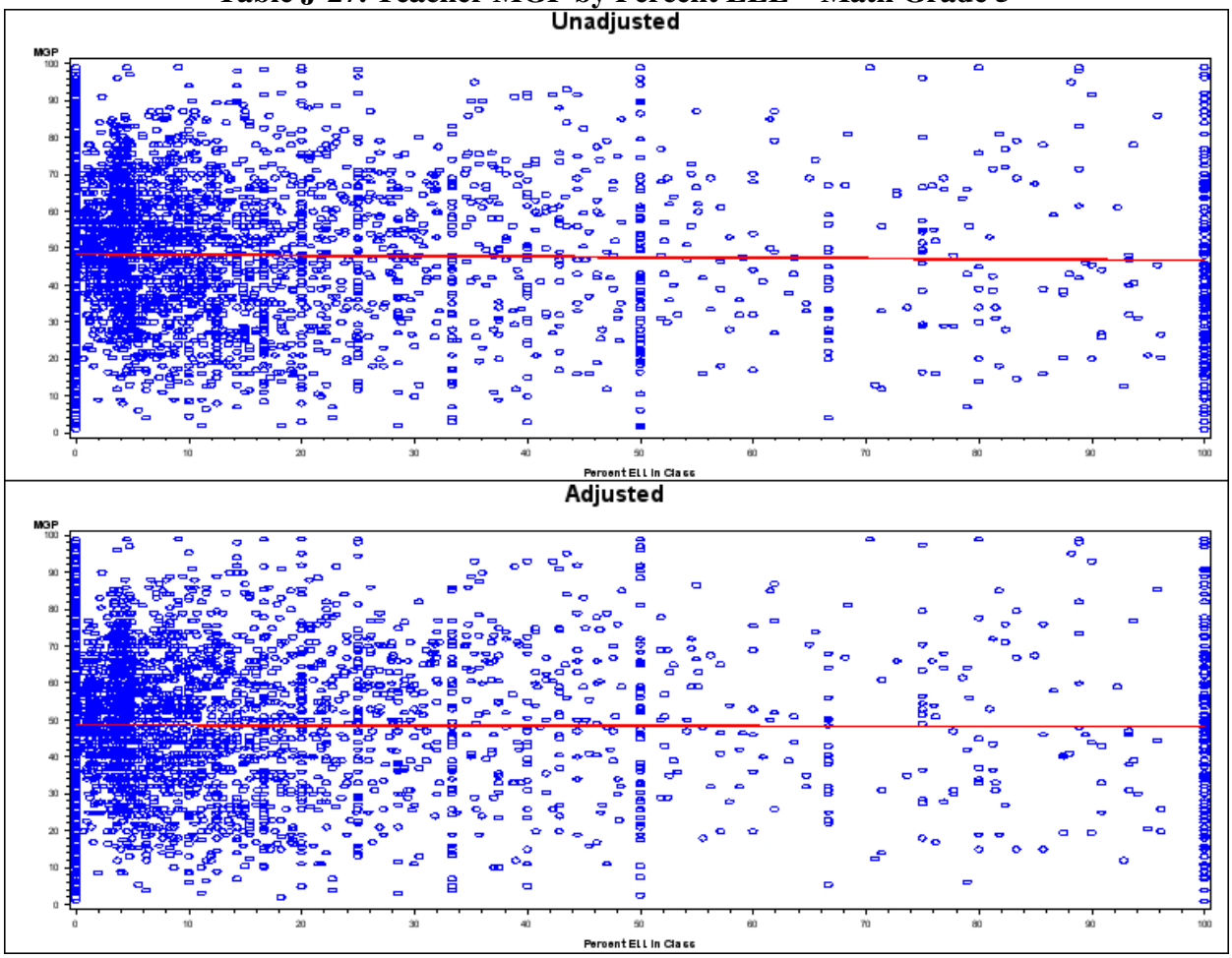


Table J-28. Teacher MGP by Percent Economically Disadvantaged—Math Grade 5

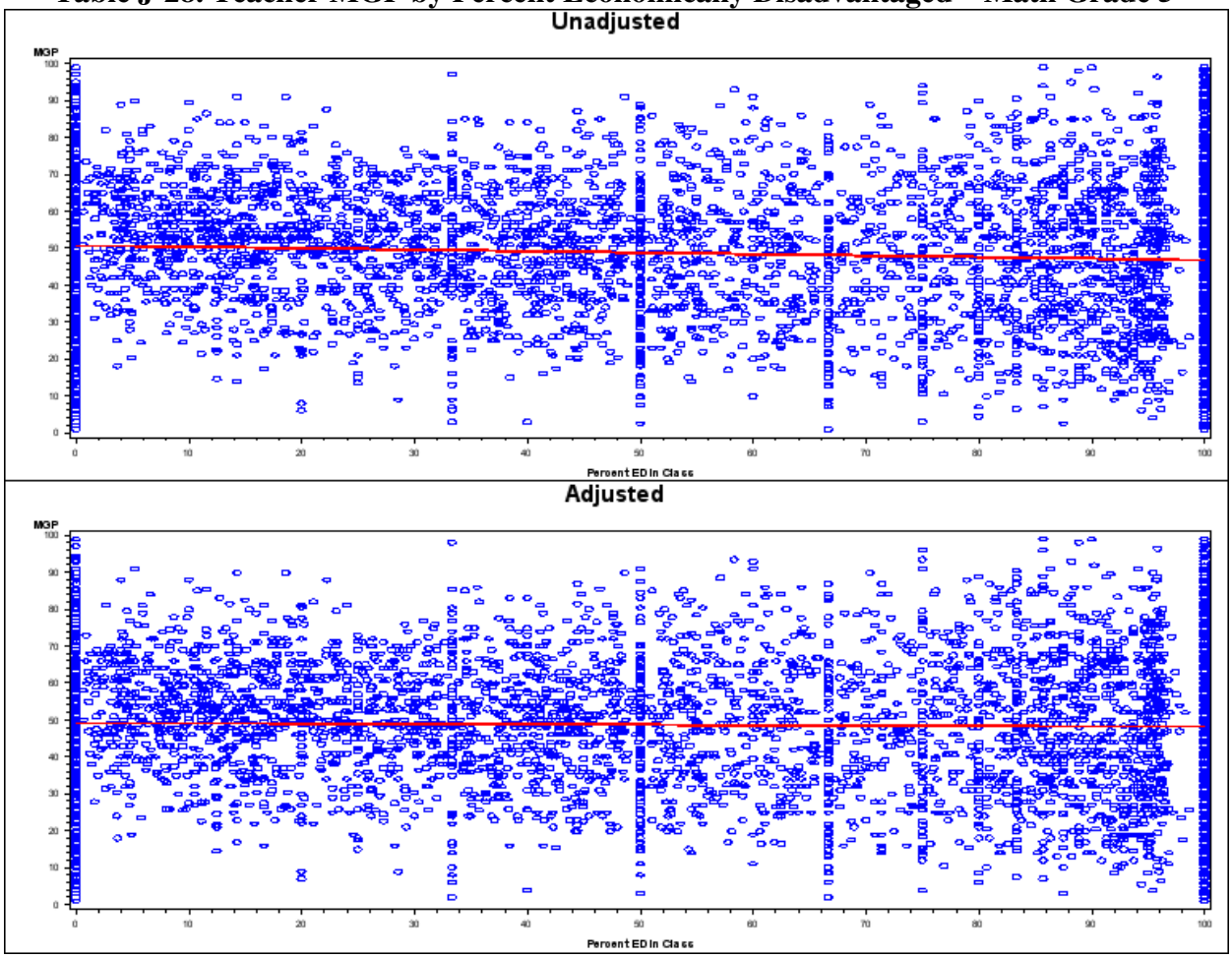


Table J-29. Teacher MGP by Students' Mean Prior Scale Score—Math Grade 6

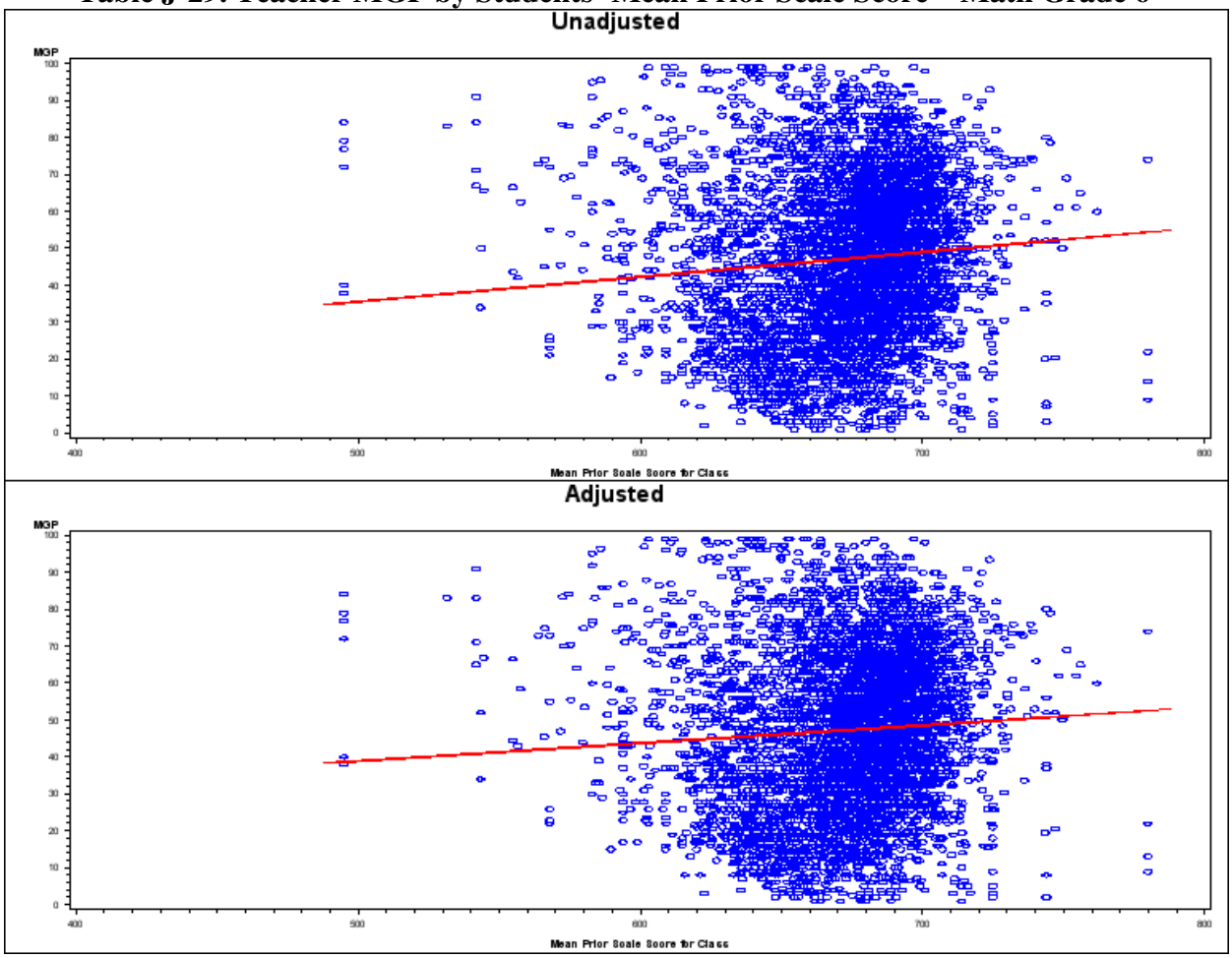


Table J-30. Teacher MGP by Percent SWD—Math Grade 6

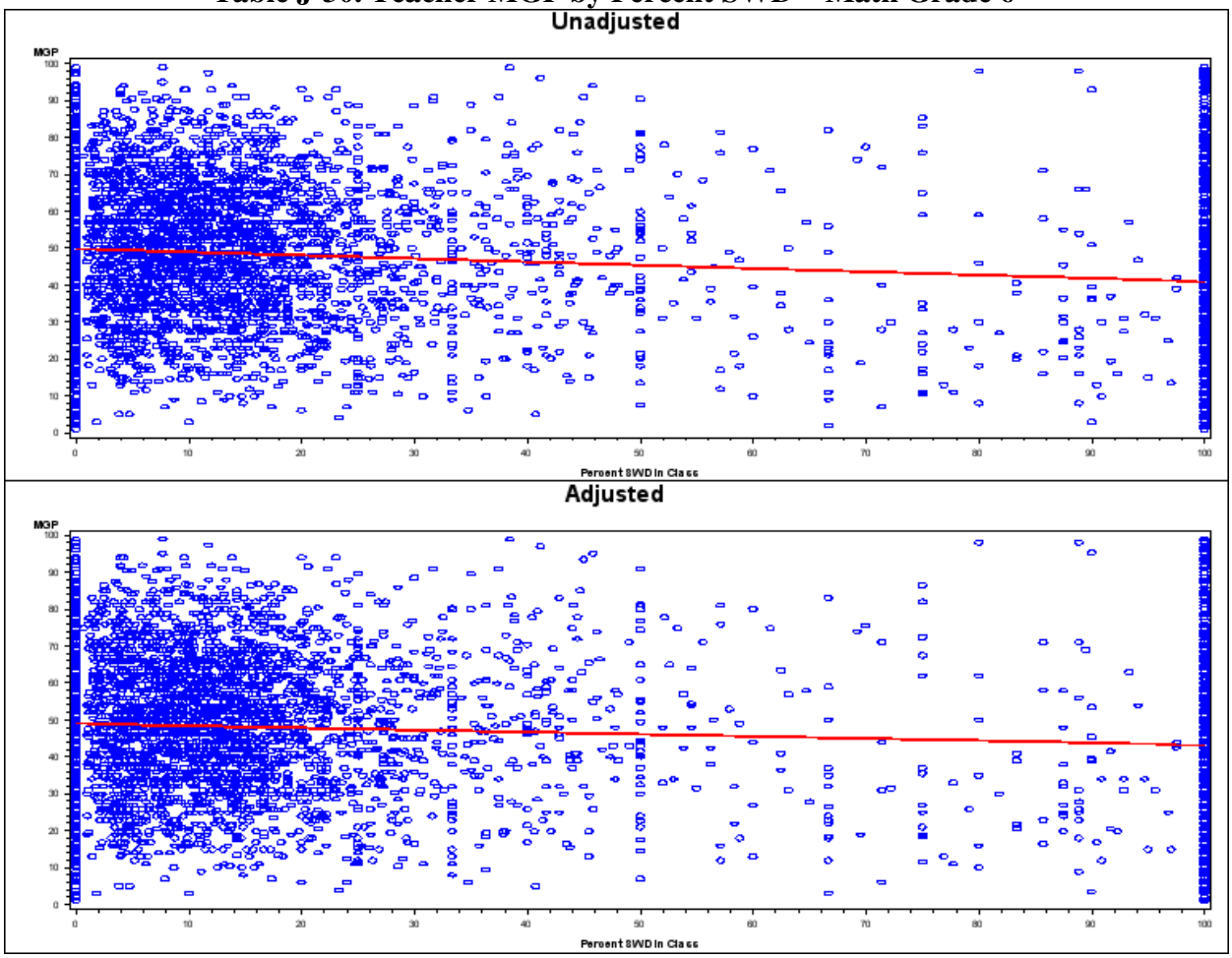


Table J-31. Teacher MGP by Percent ELL—Math Grade 6

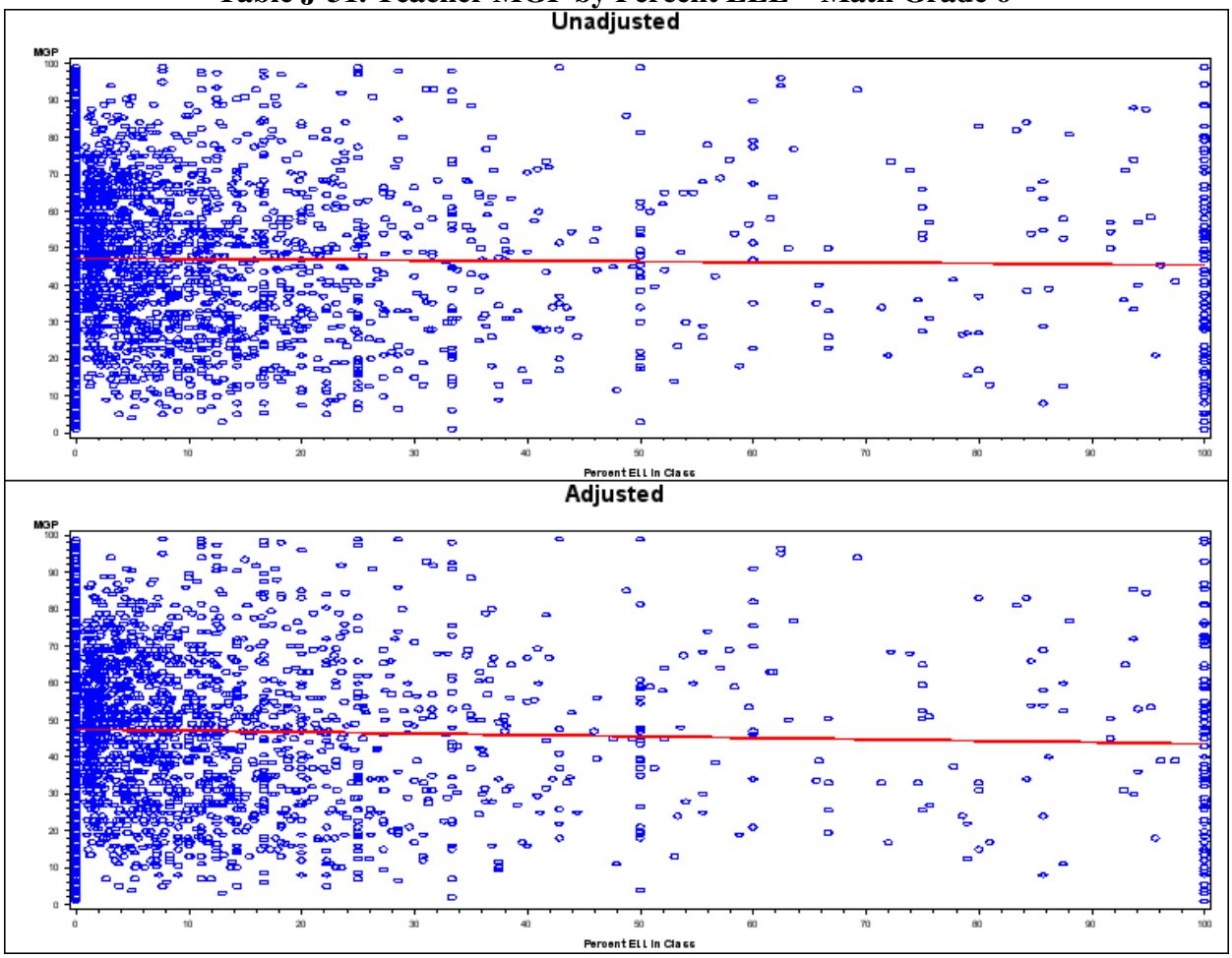


Table J-32. Teacher MGP by Percent Economically Disadvantaged—Math Grade 6

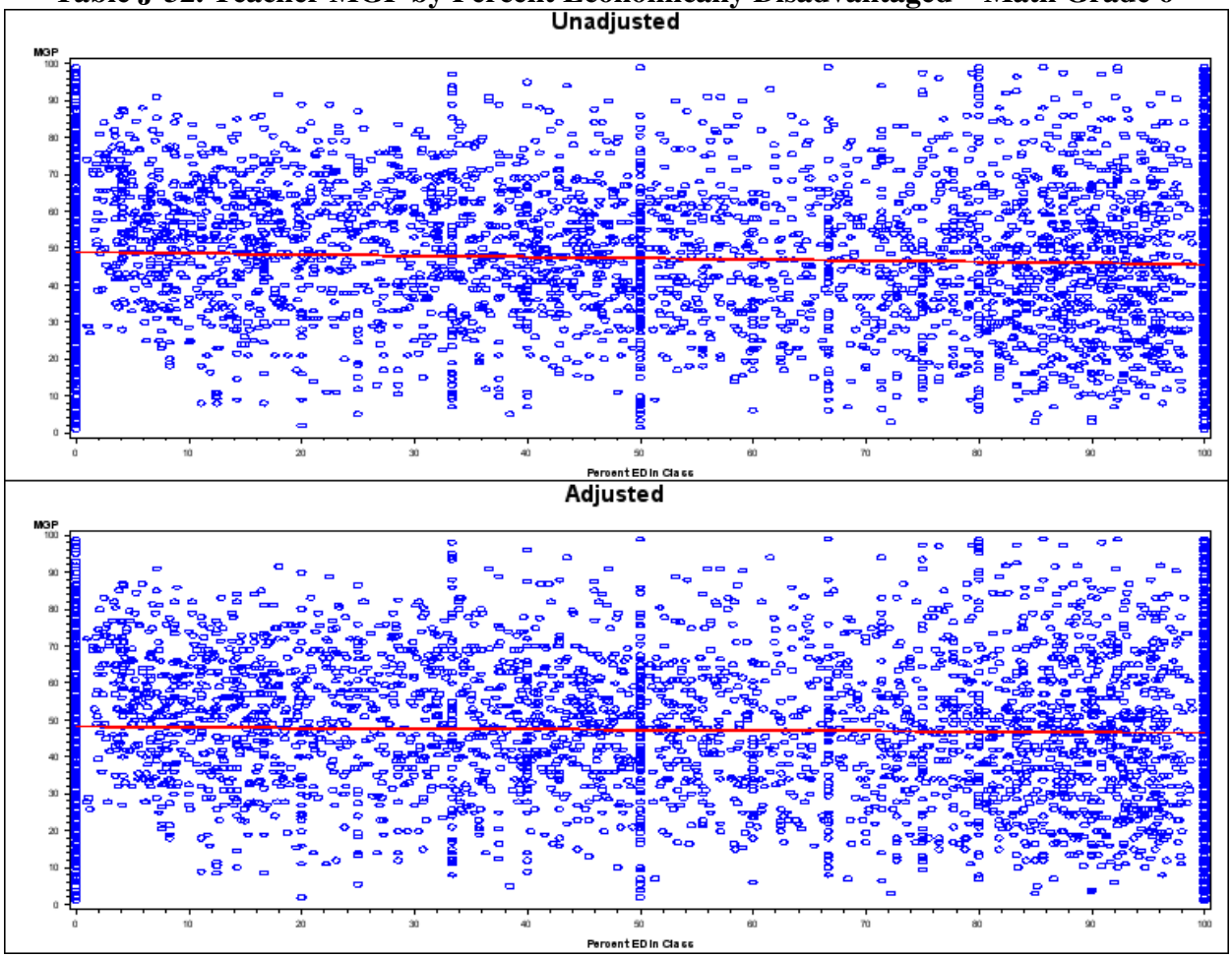


Table J-33. Teacher MGP by Students' Mean Prior Scale Score—Math Grade 7

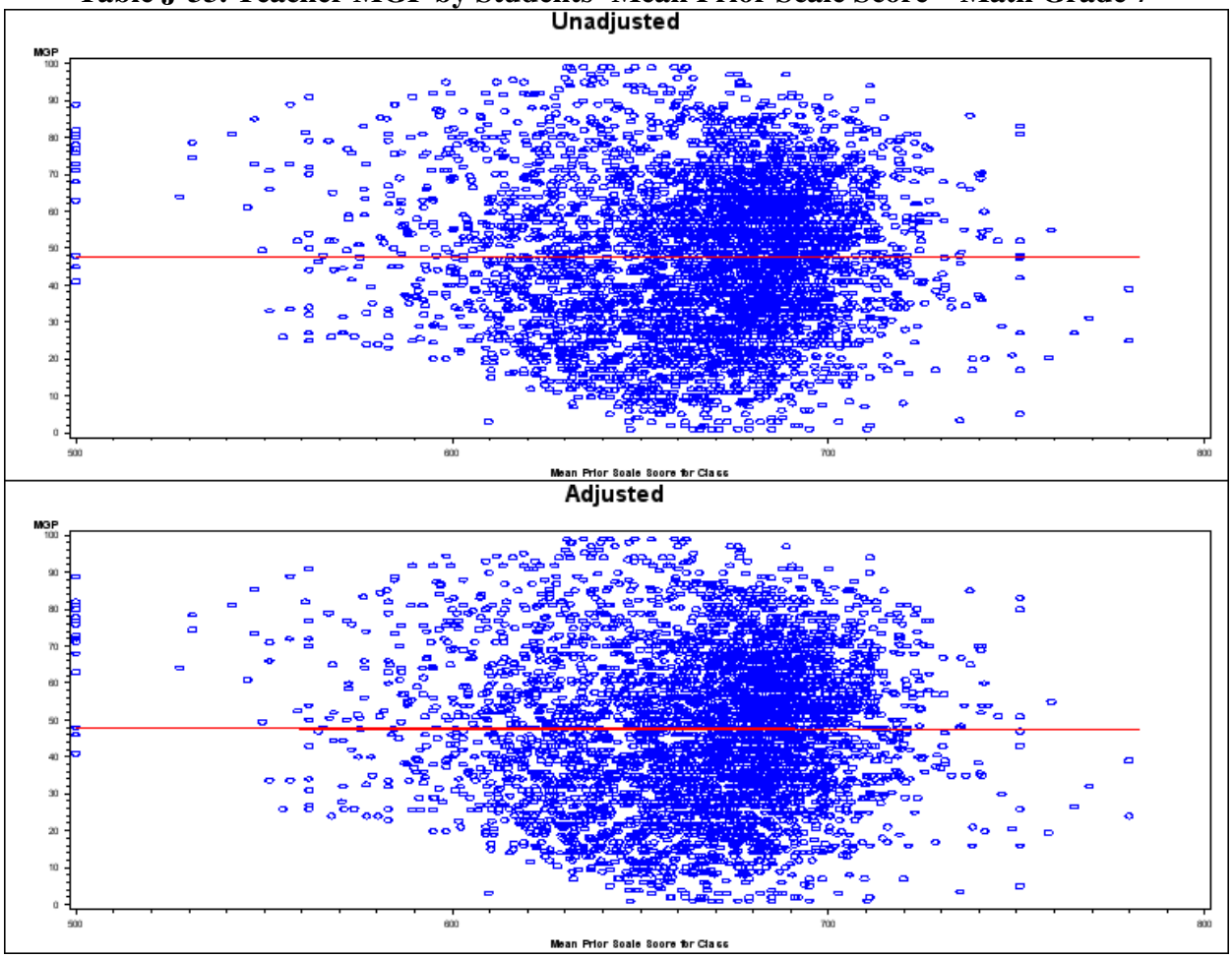


Table J-34. Teacher MGP by Percent SWD—Math Grade 7

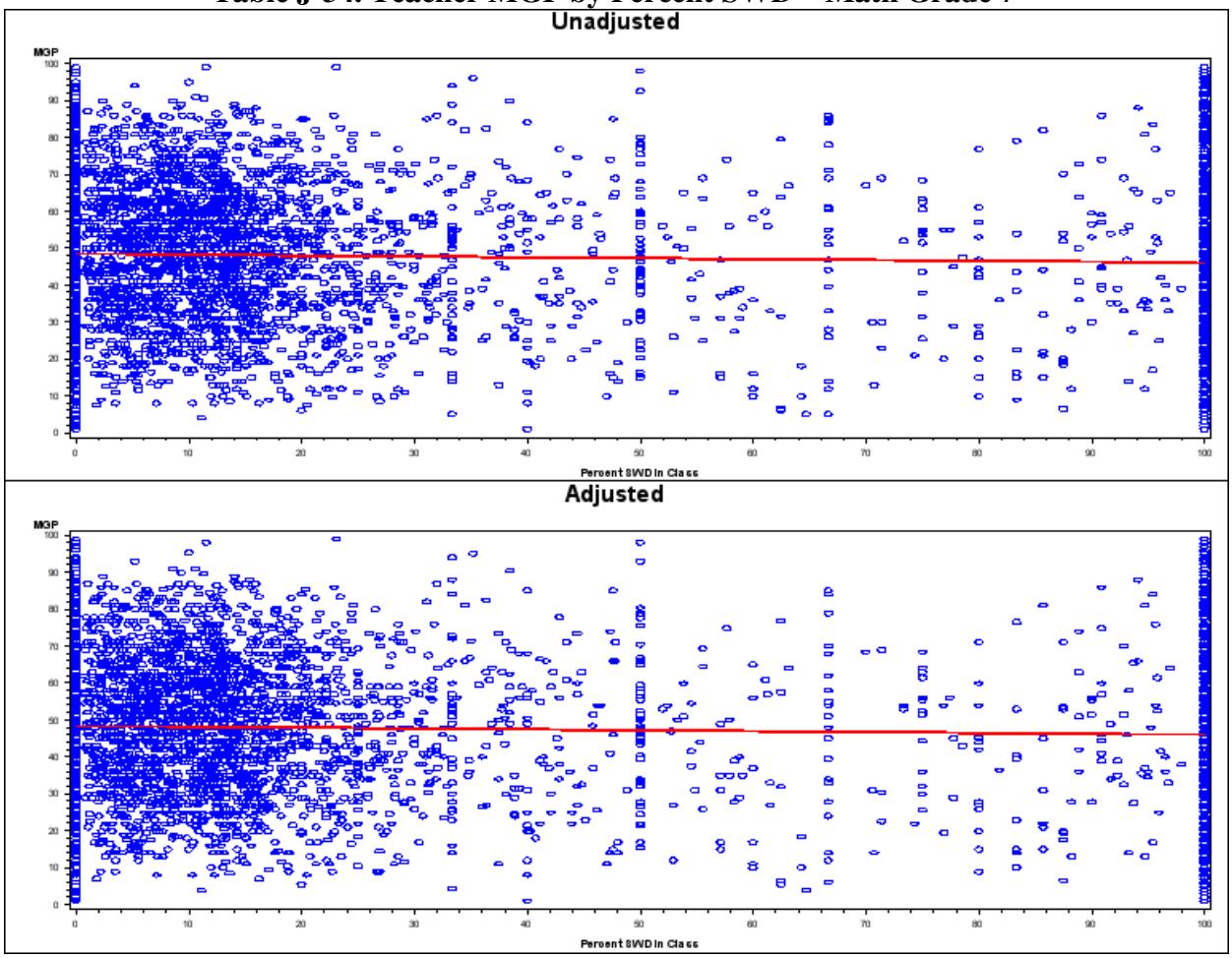


Table J-35. Teacher MGP by Percent ELL—Math Grade 7

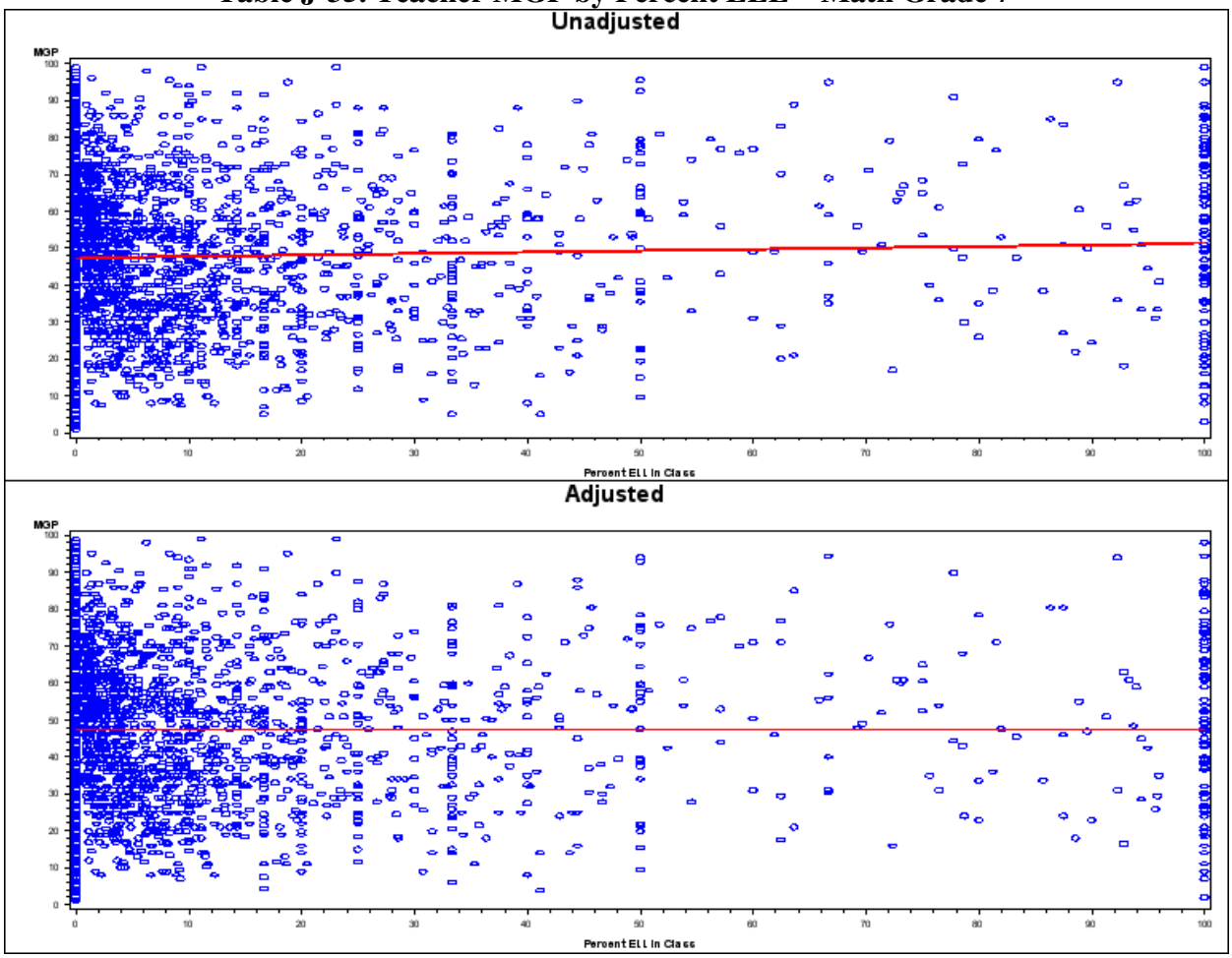


Table J-36. Teacher MGP by Percent Economically Disadvantaged—Math Grade 7

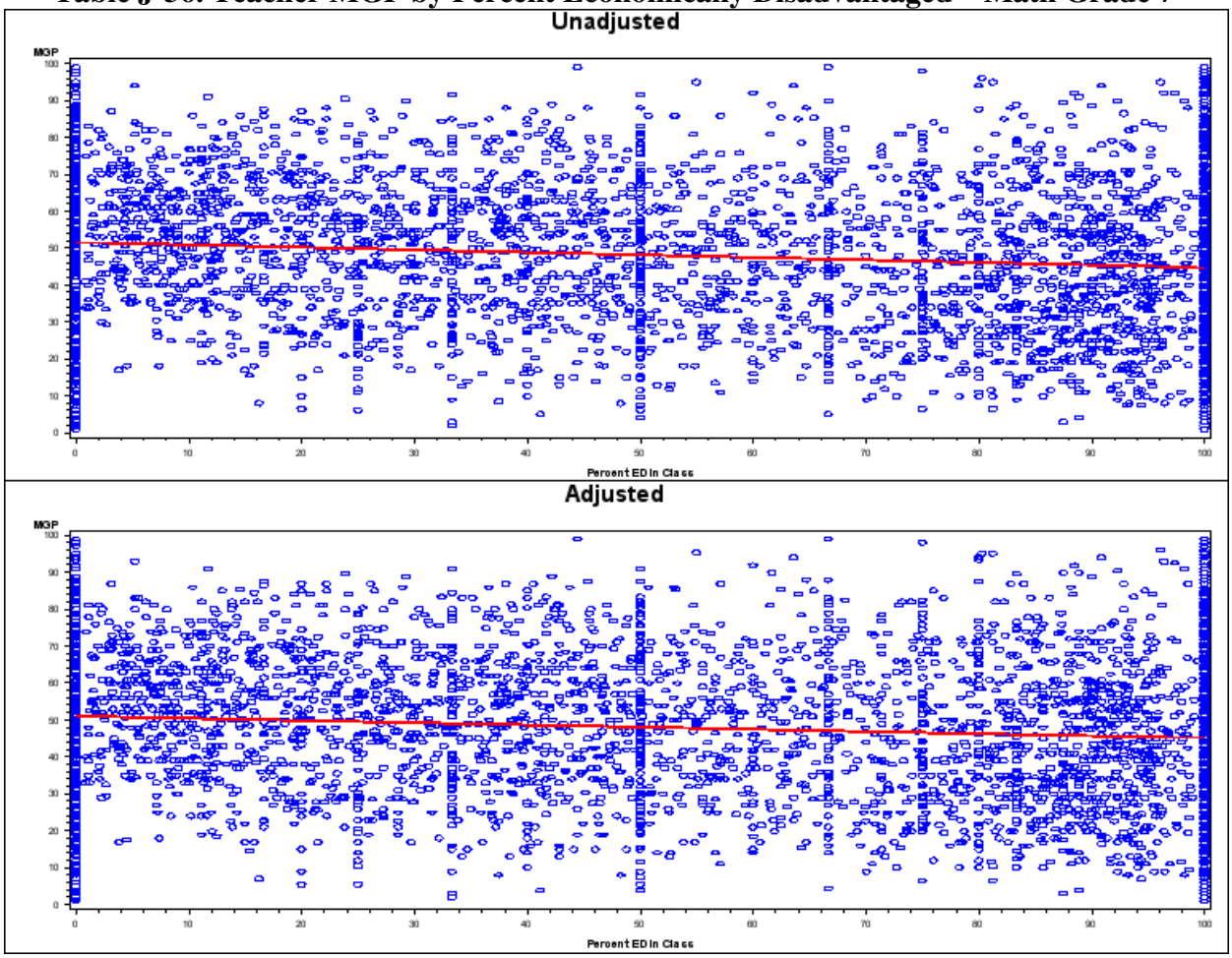


Table J-37. Teacher MGP by Students' Mean Prior Scale Score—Math Grade 8

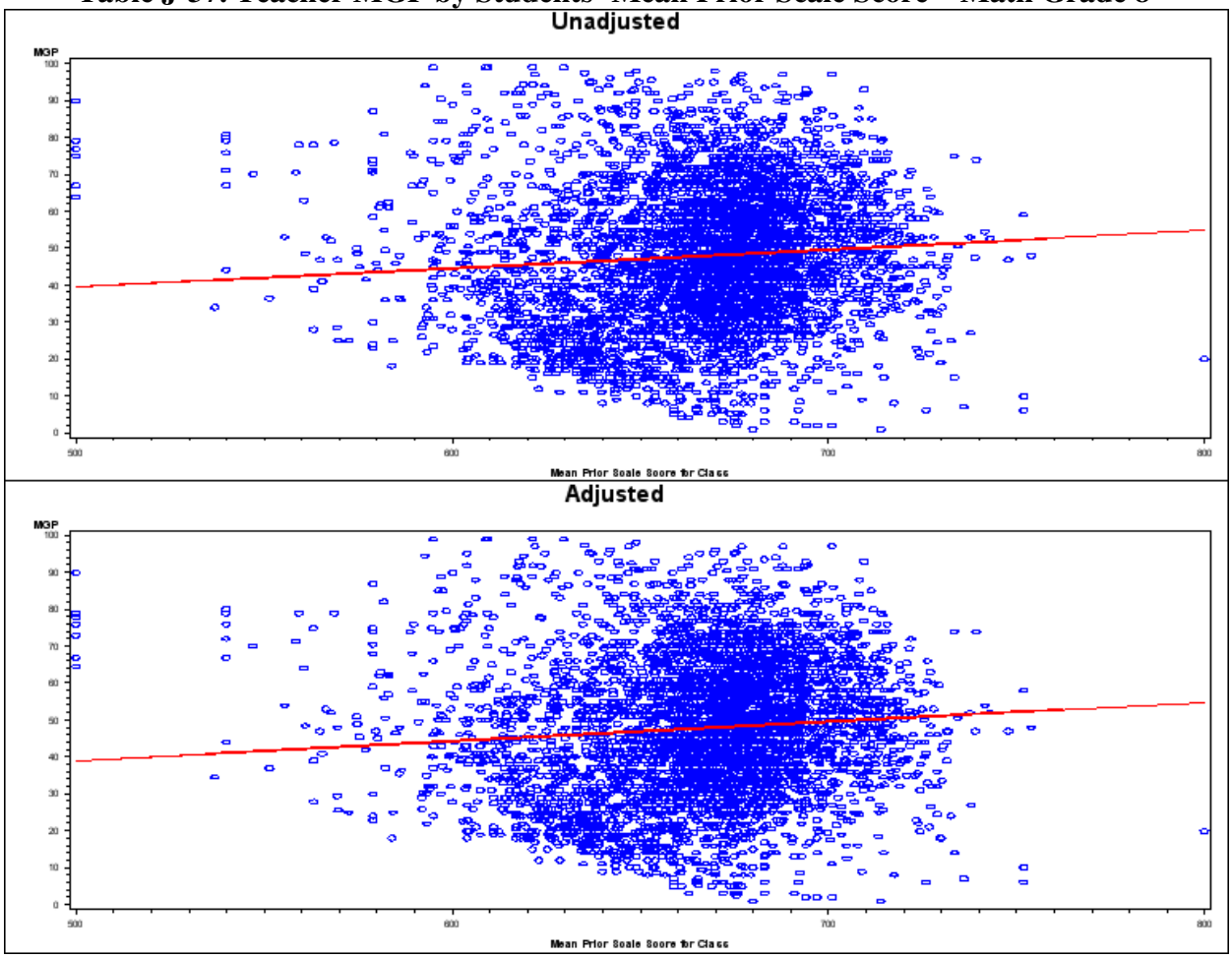


Table J-38. Teacher MGP by Percent SWD—Math Grade 8

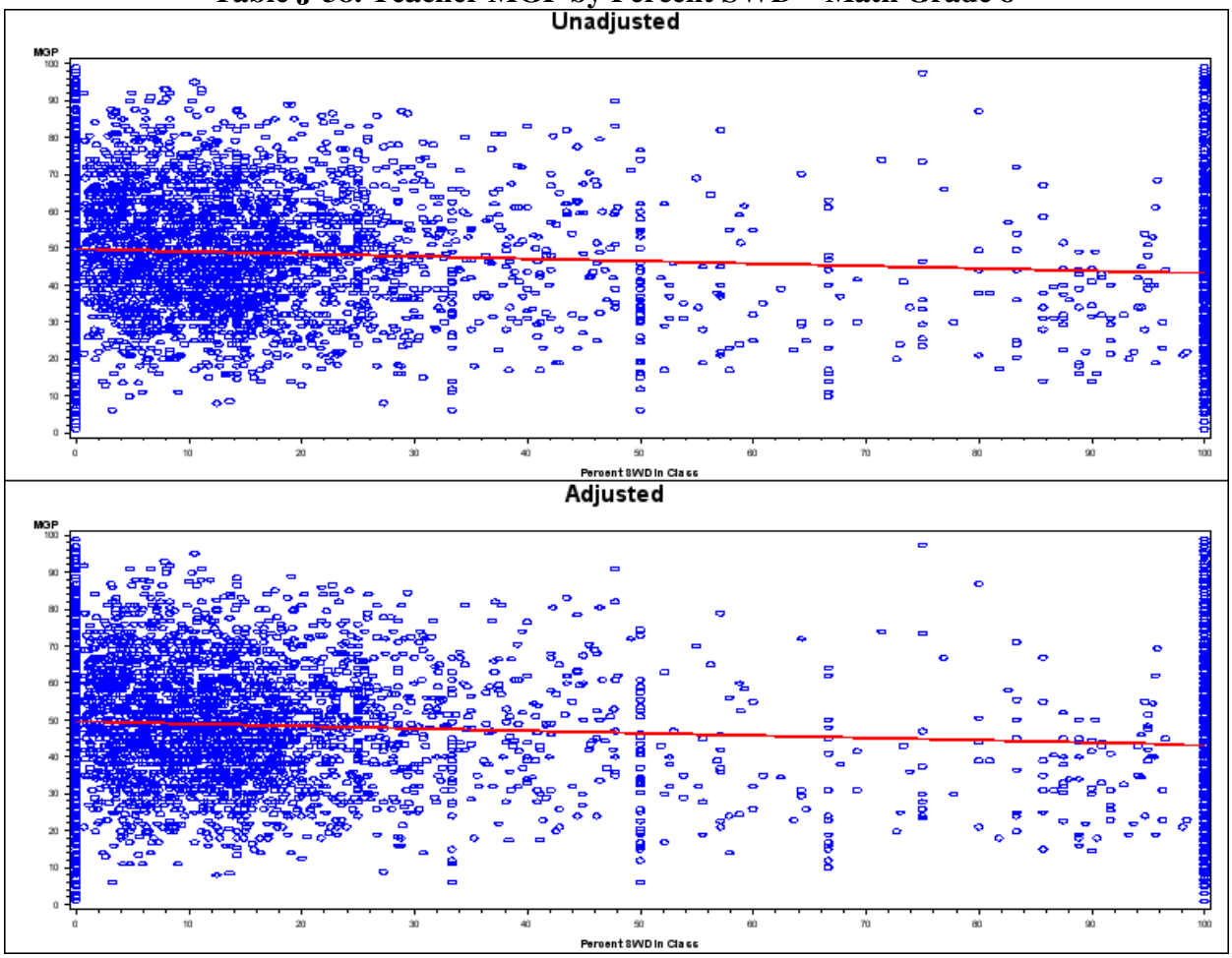


Table J-39. Teacher MGP by Percent ELL—Math Grade 8

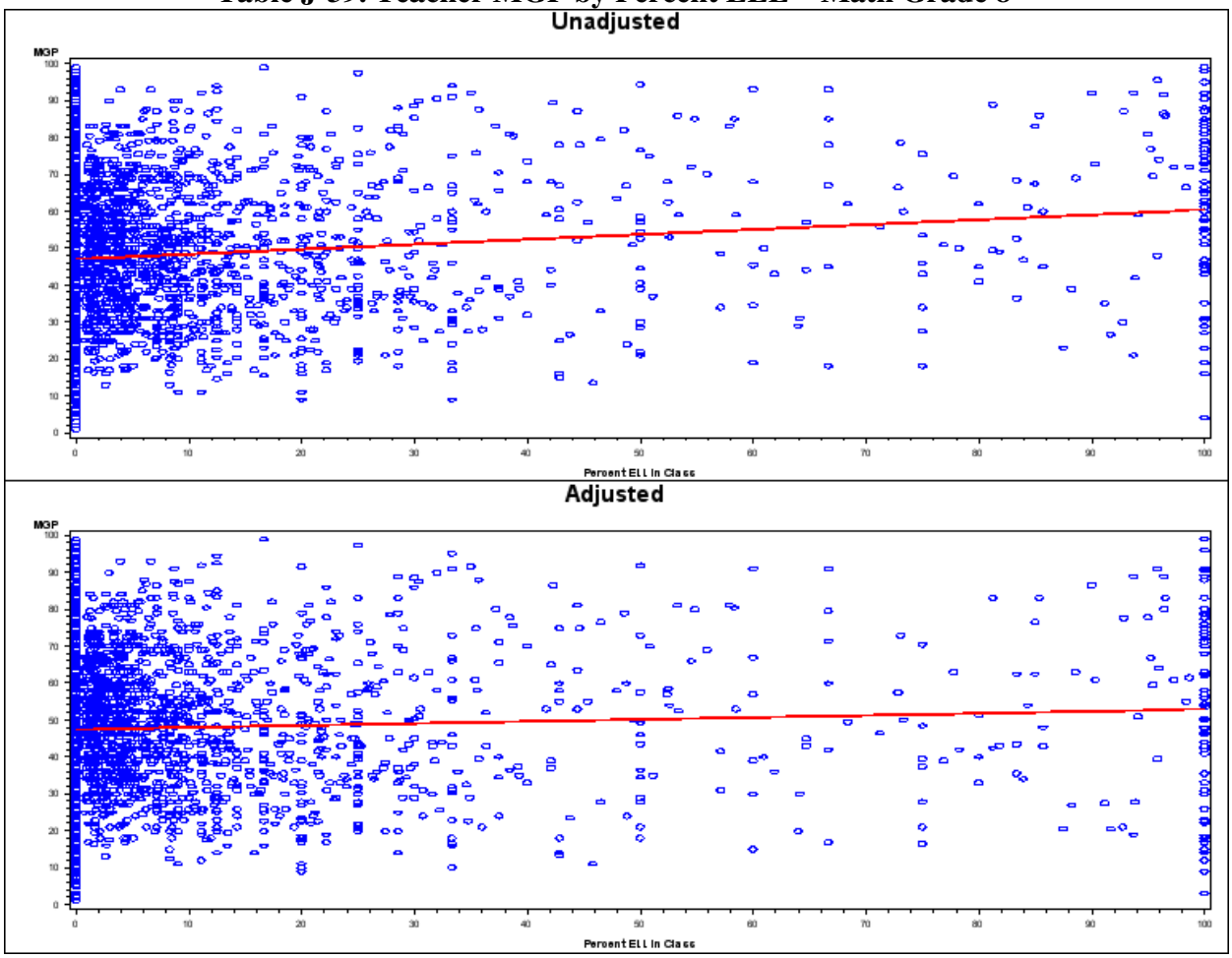


Table J-40. Teacher MGP by Percent Economically Disadvantaged—Math Grade 8

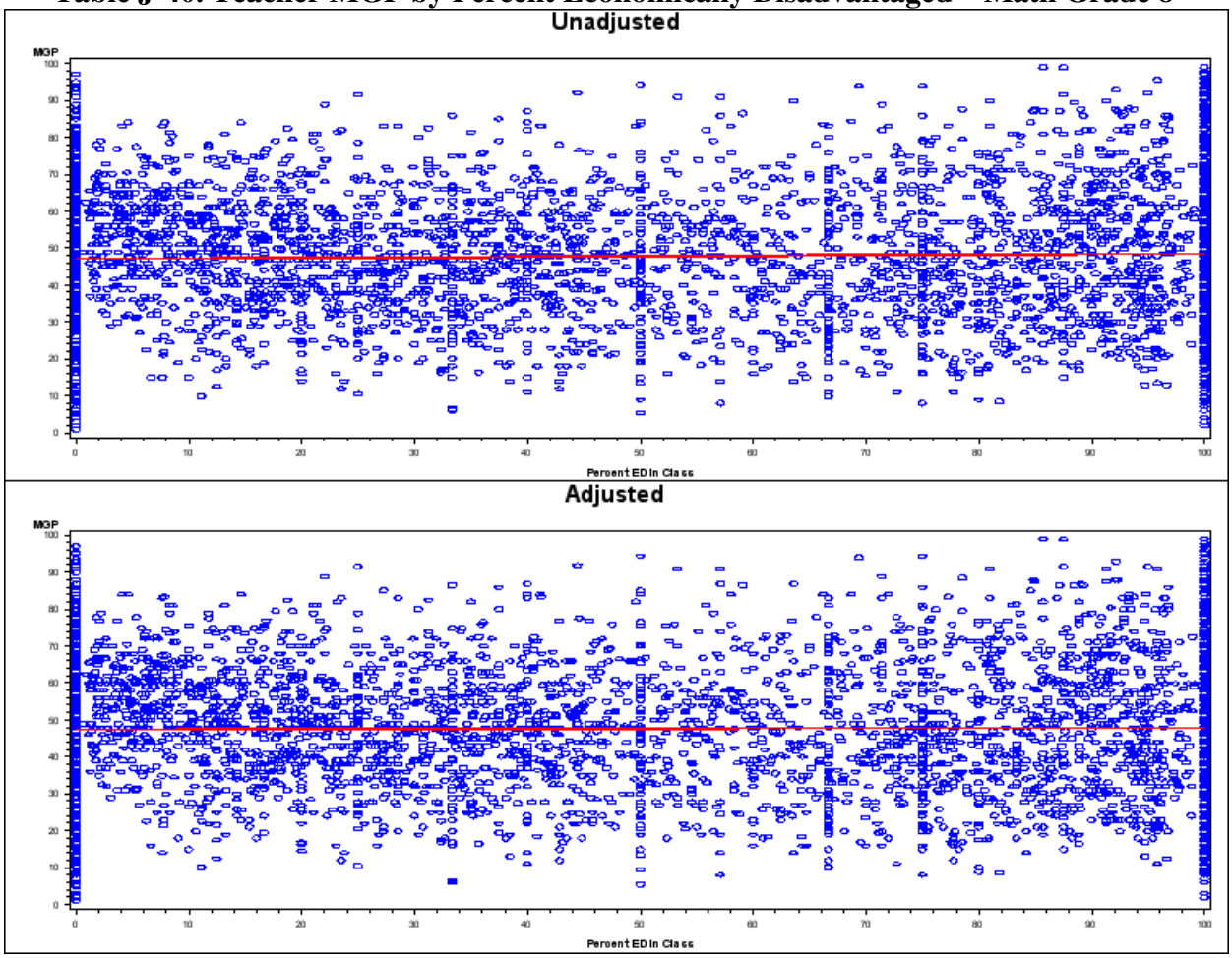


Table J-41. School MGP by Students' Mean Prior Scale Score—ELA Combined Grades

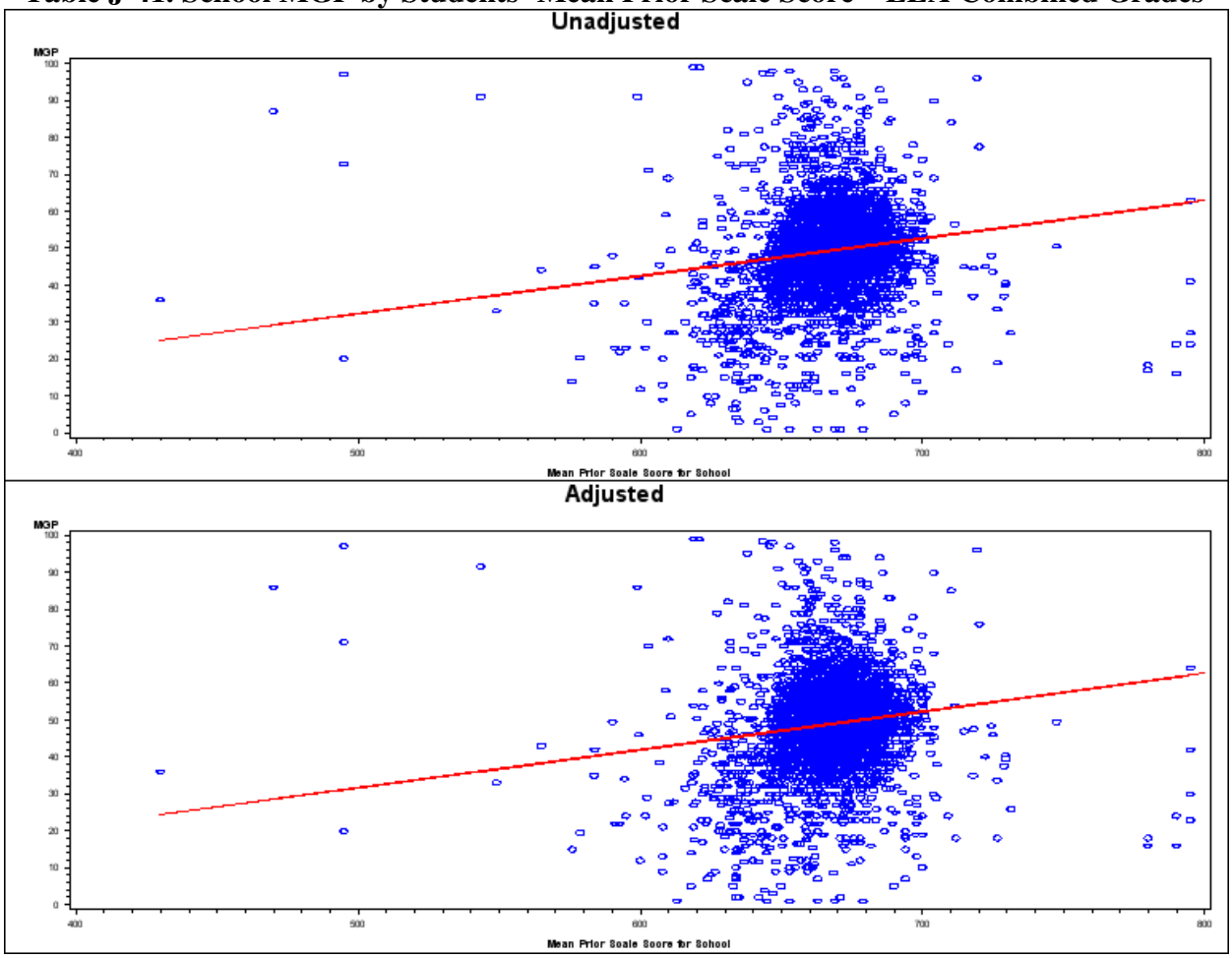


Table J-42. School MGP by Percent SWD—ELA Combined Grades

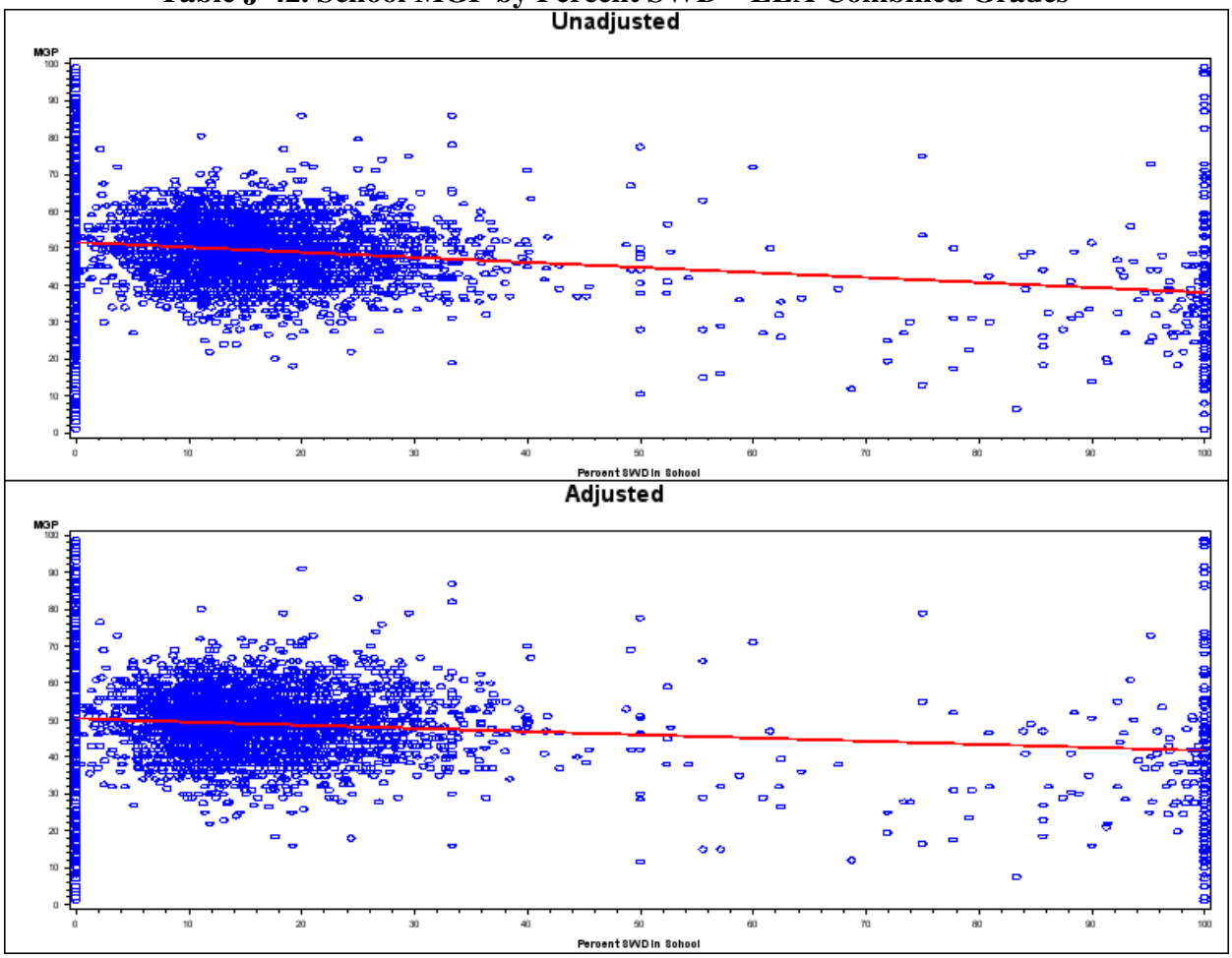


Table J-43. School MGP by Percent ELL—ELA Combined Grades

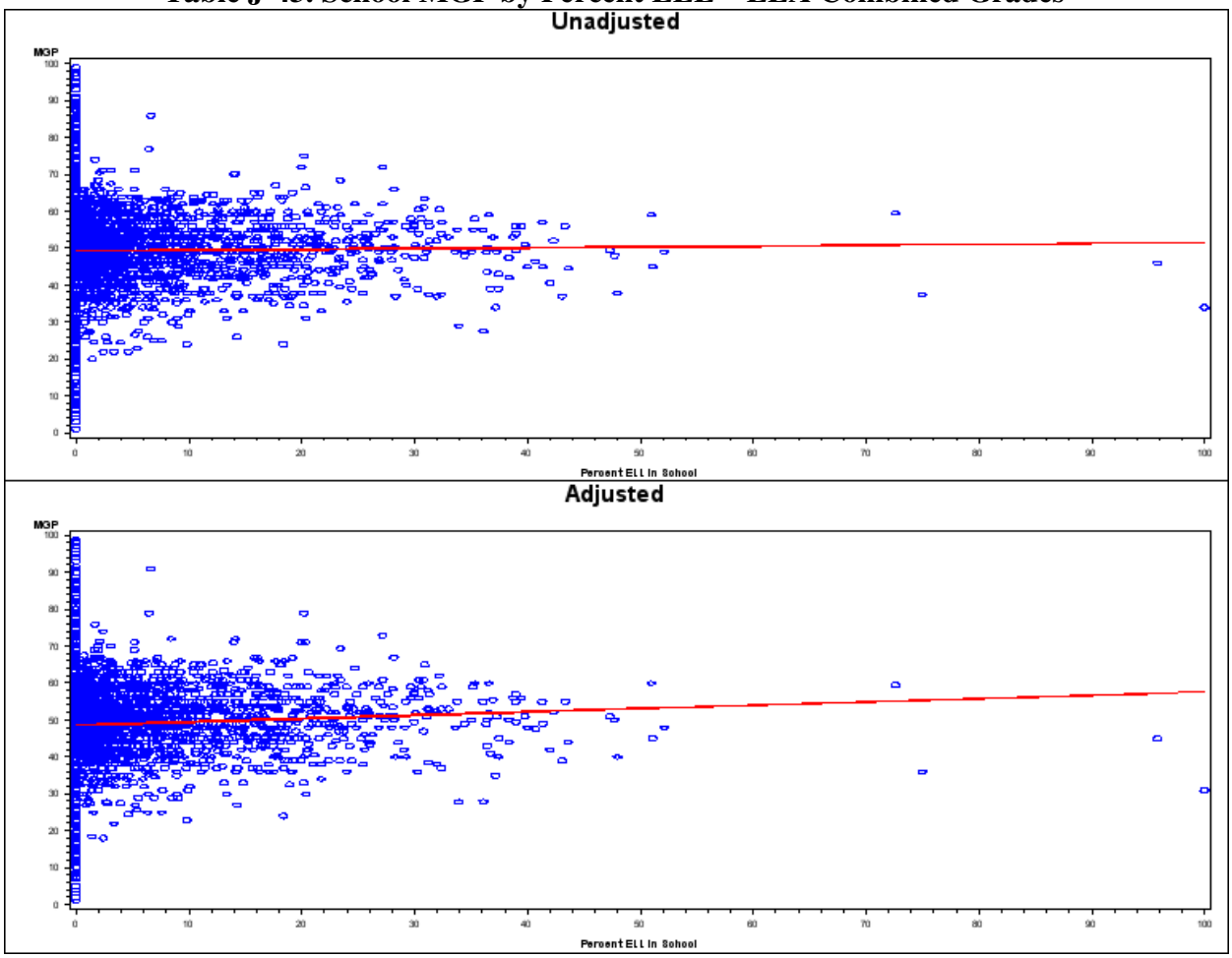


Table J-44. School MGP by Percent Economically Disadvantaged—ELA Combined Grades

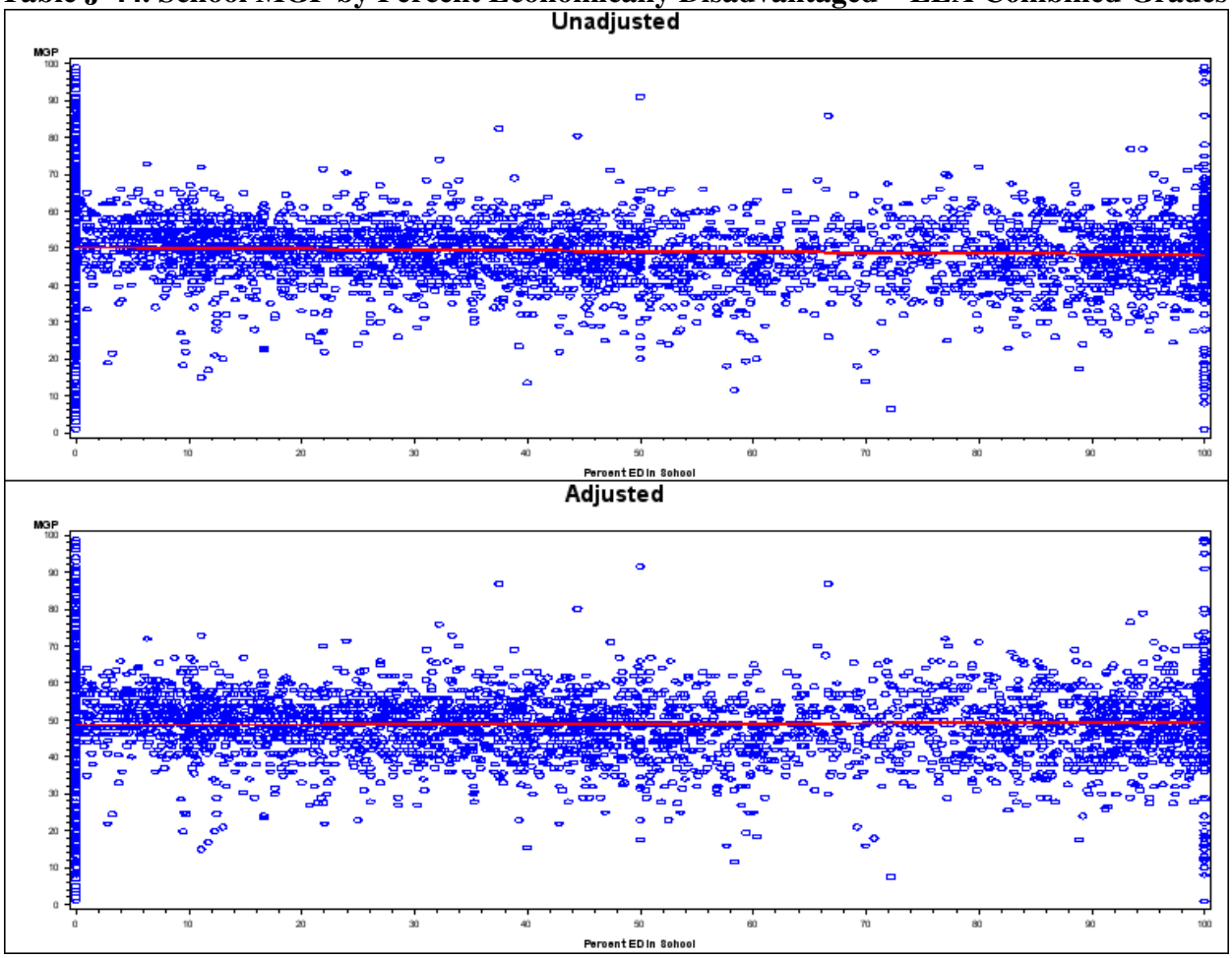


Table J-45. School MGP by Students' Mean Prior Scale Score—Math Combined Grades

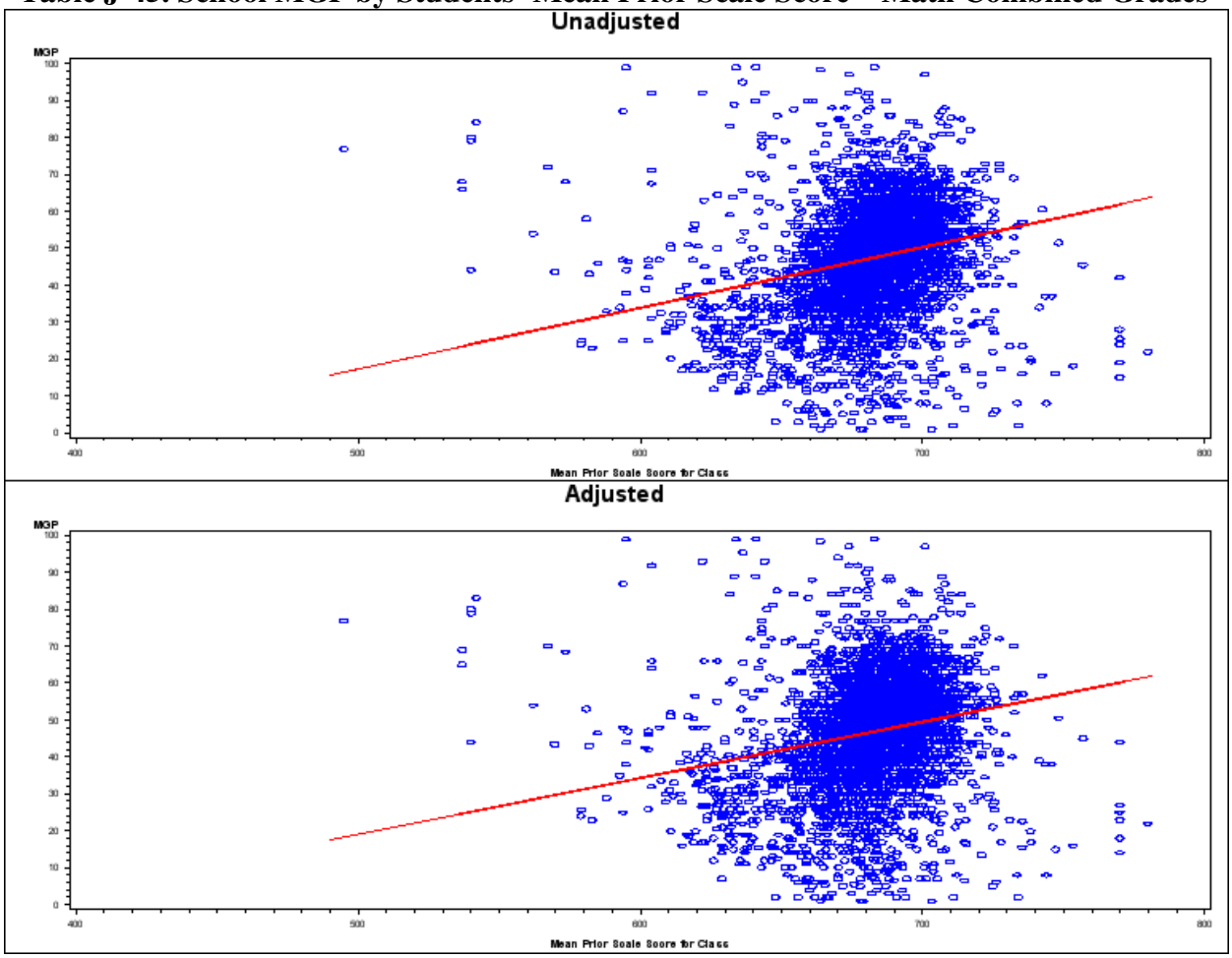


Table J-46. School MGP by Percent SWD—Math Combined Grades

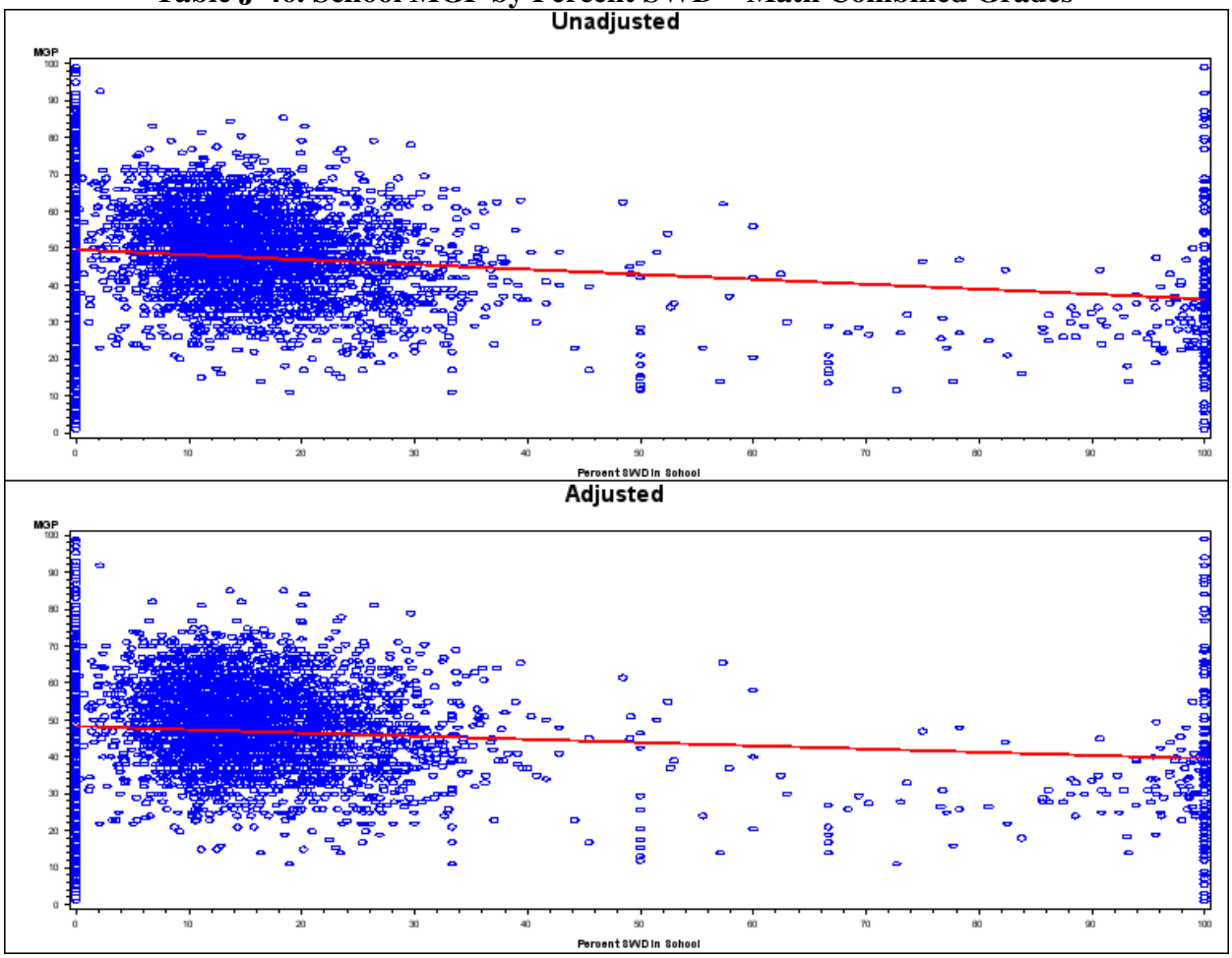


Table J-47. School MGP by Percent ELL—Math Combined Grades

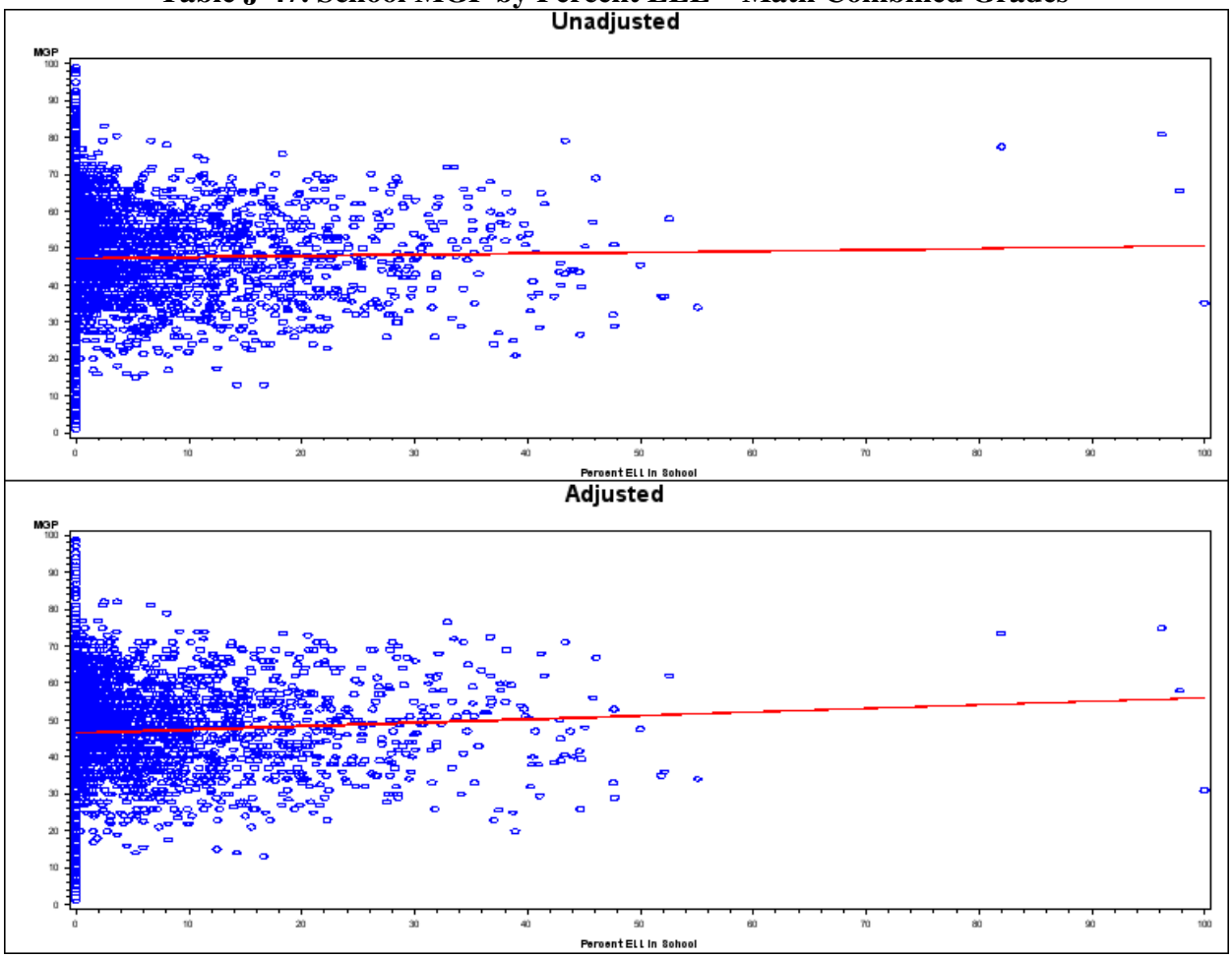


Table J-48. School MGP by Percent Economically Disadvantaged—Math Combined Grades

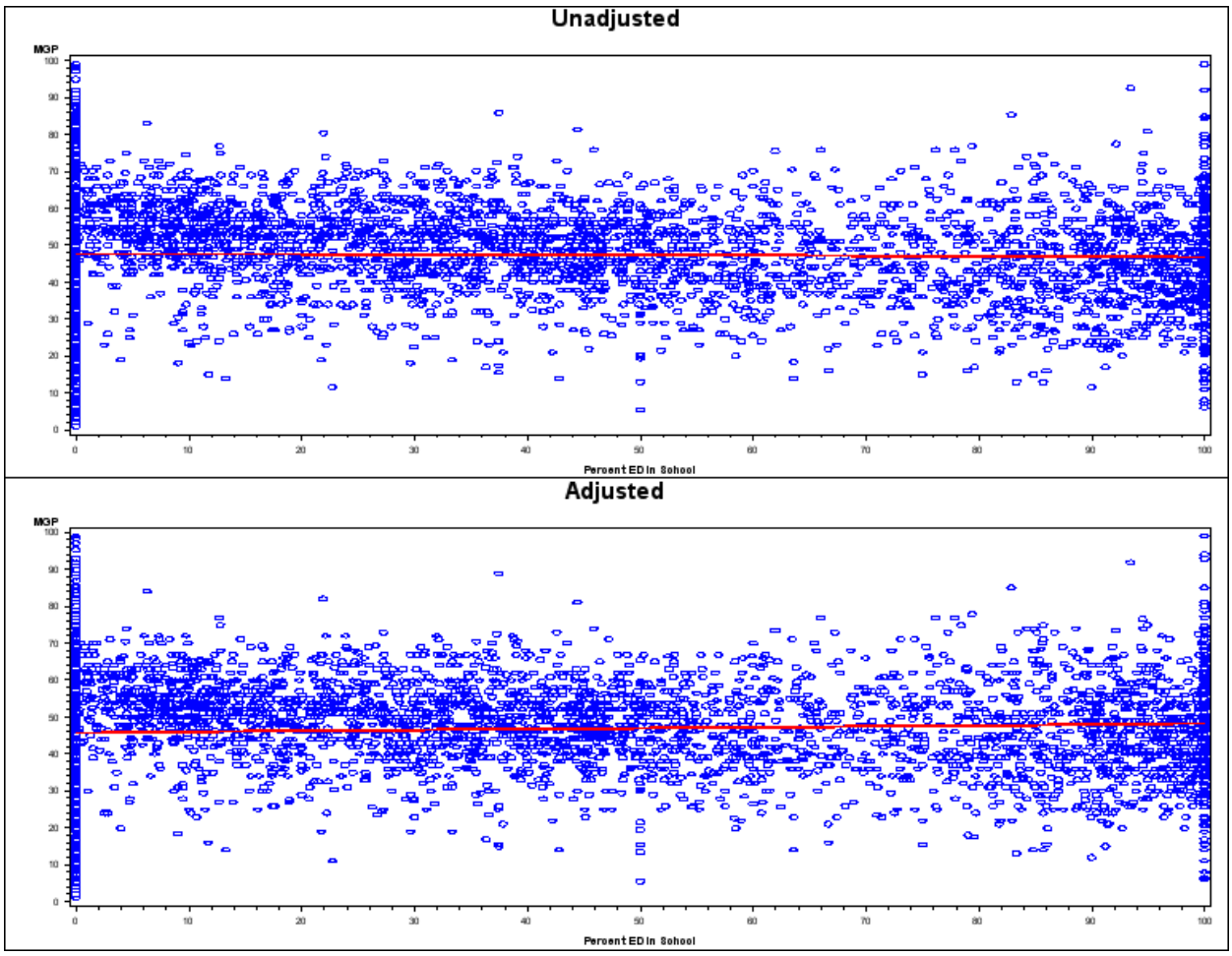


Table J-49. School MGP by Students' Mean Prior ELA Scale Score—Combined Grades and Subjects

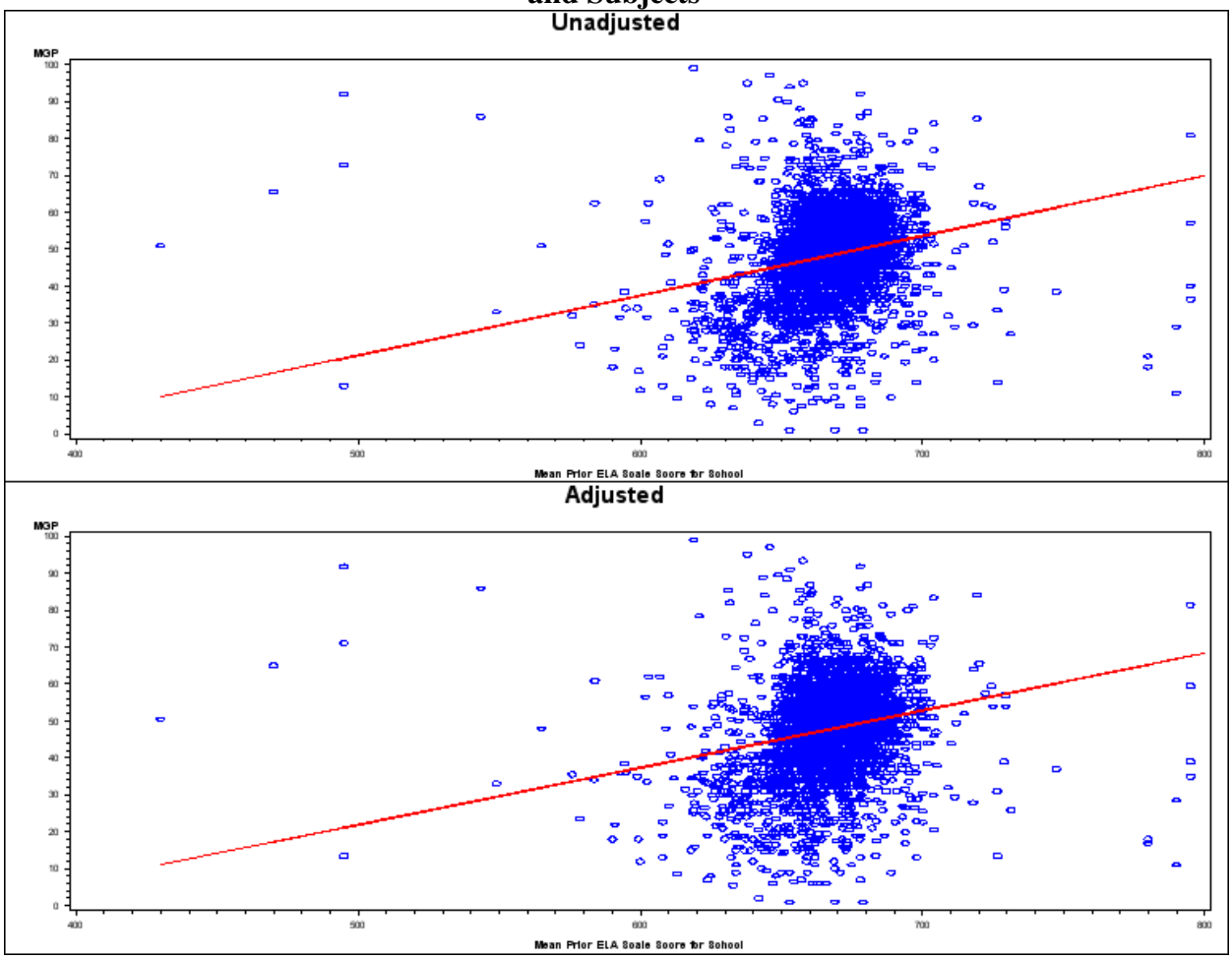


Table J-50. School MGP by Students' Mean Prior Math Scale Score—Combined Grades and Subjects

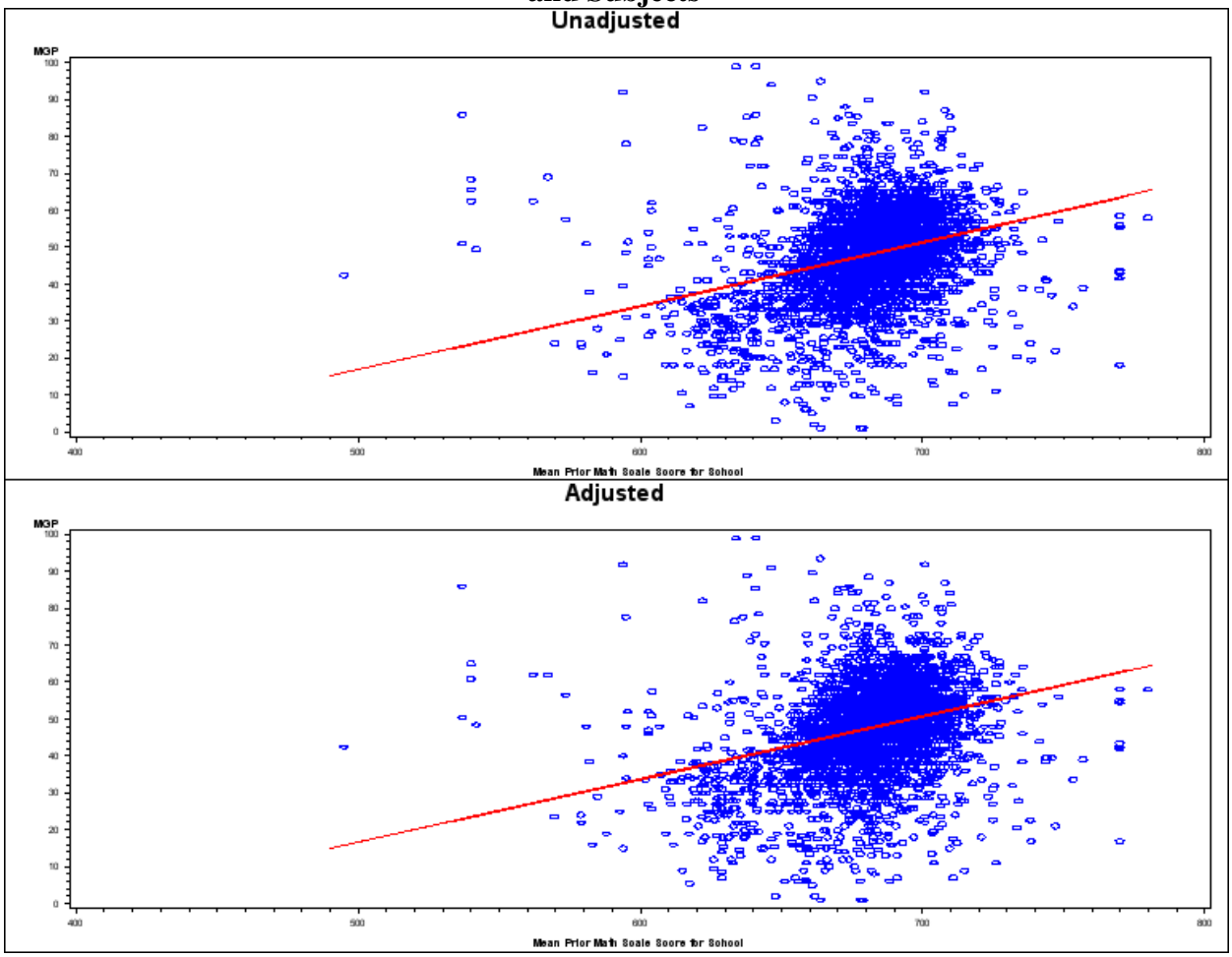


Table J-51. School MGP by Percent SWD—Combined Grades and Subjects

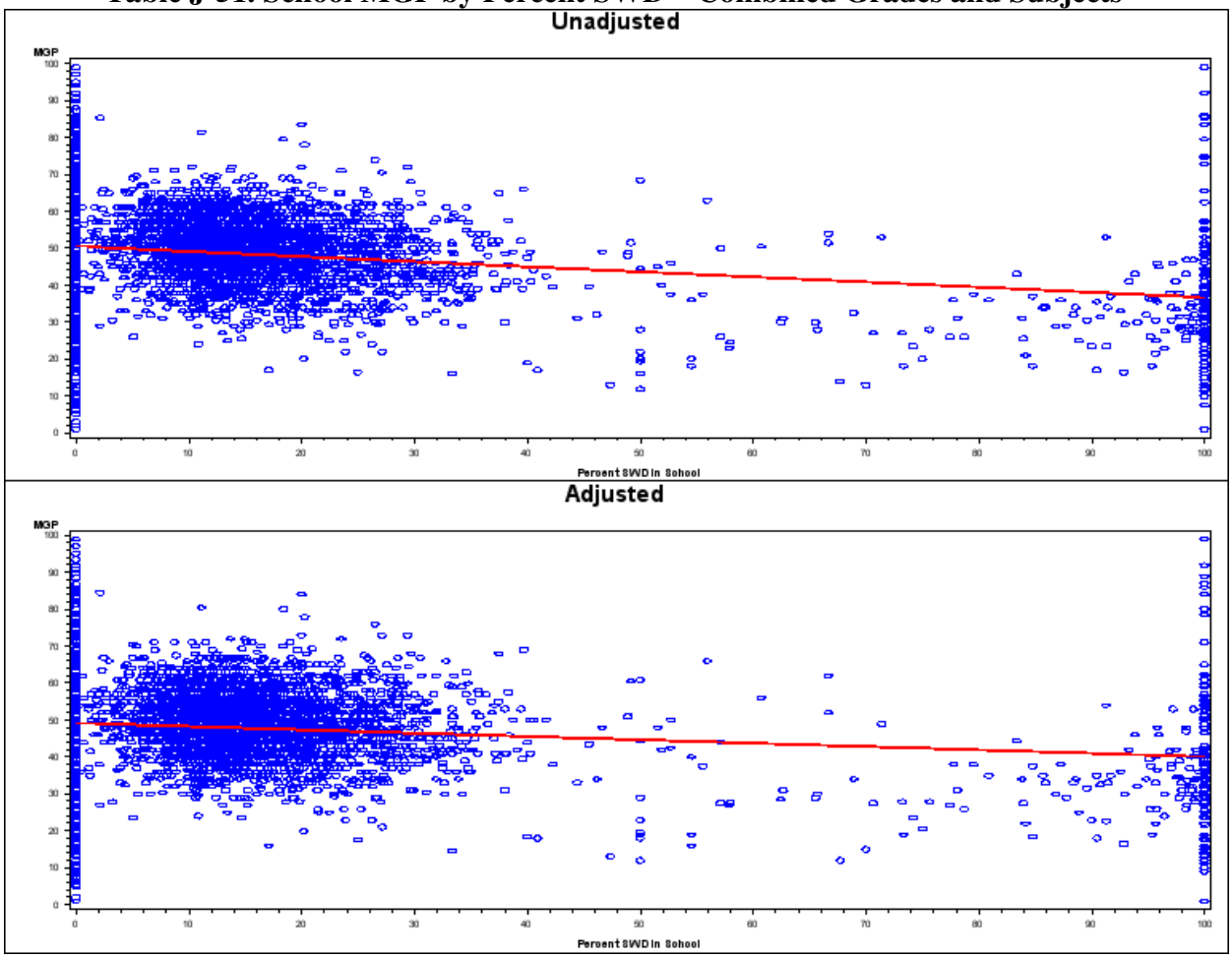


Table J-52. School MGP by Percent ELL—Combined Grades and Subjects

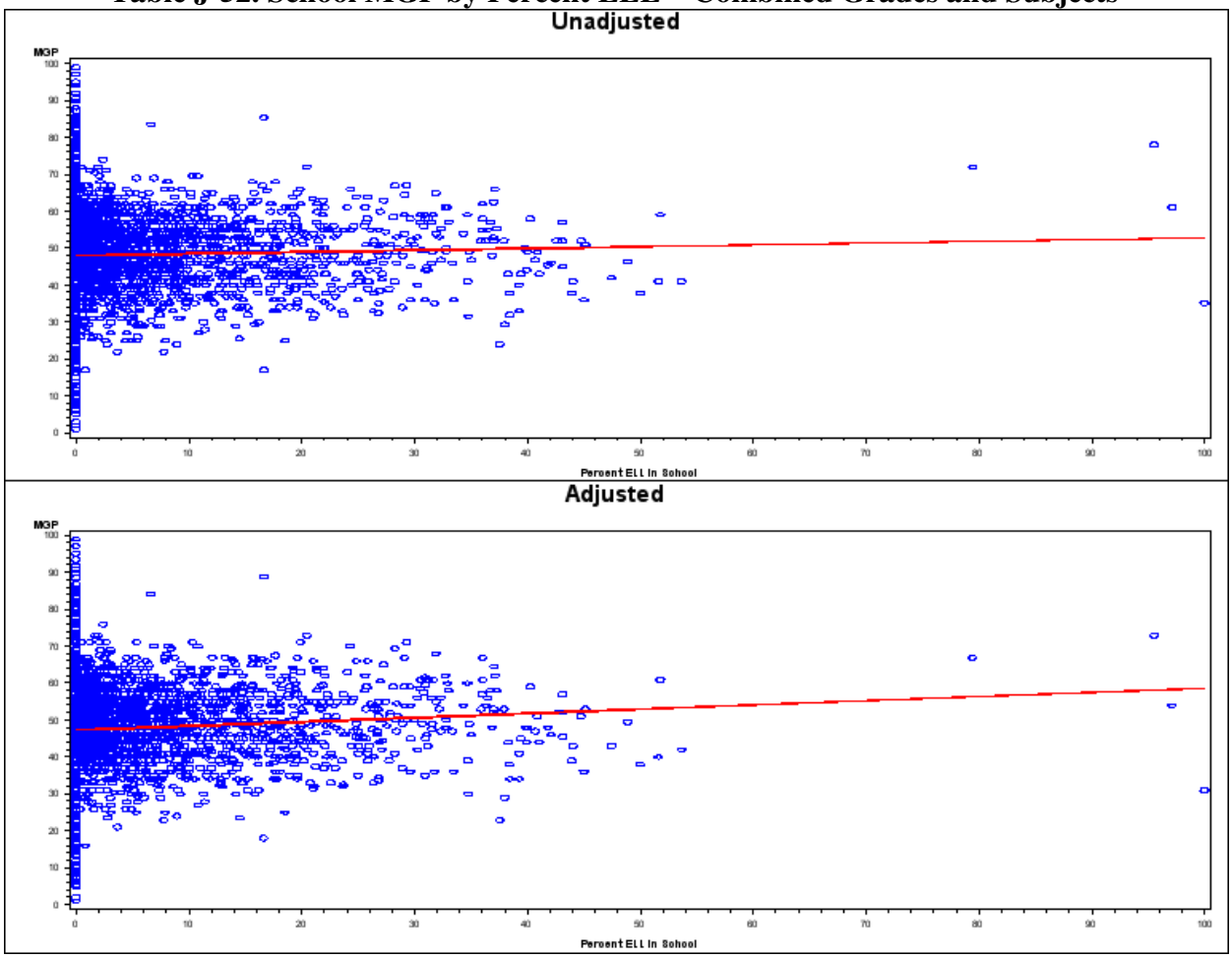
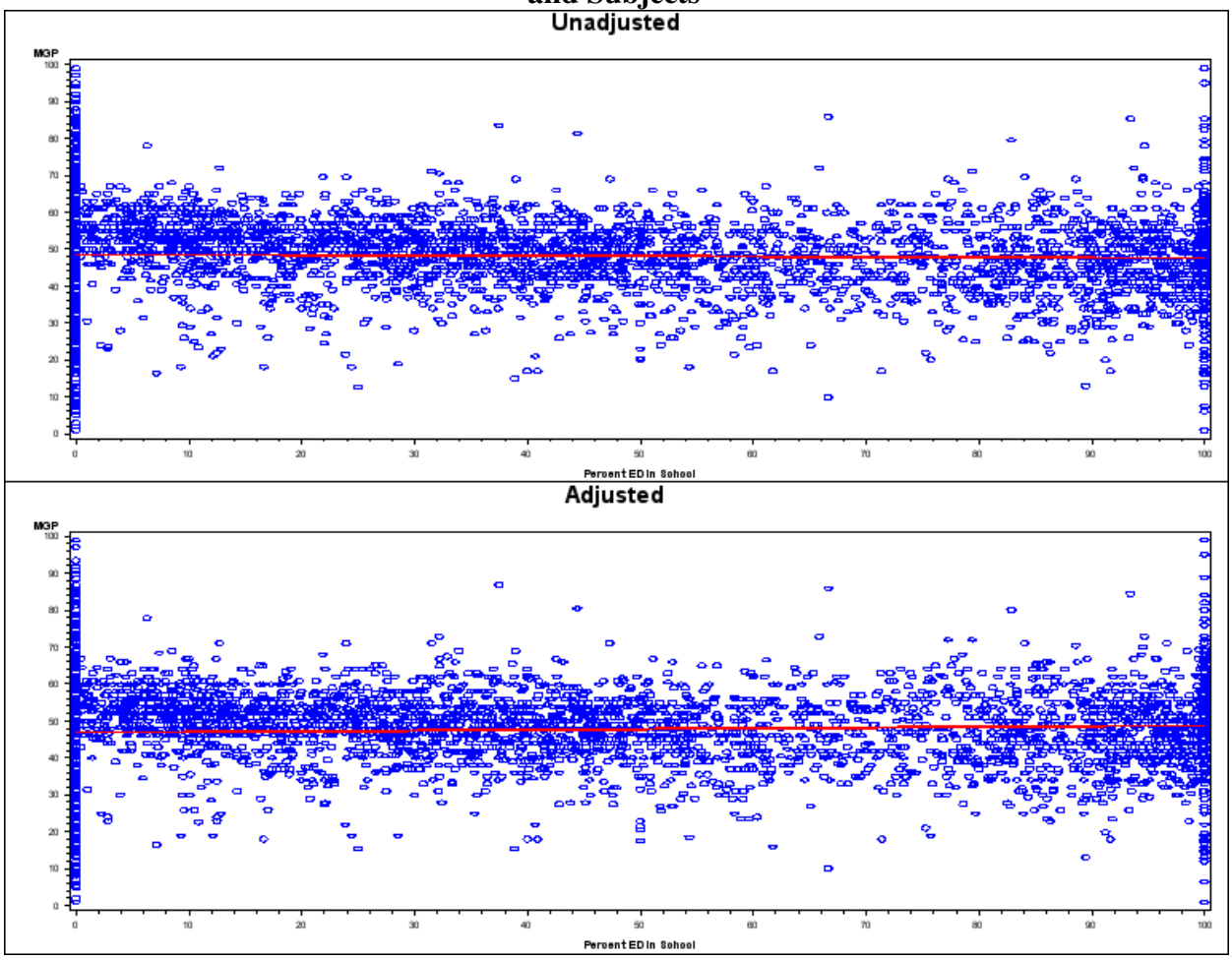


Table J-53. School MGP by Percent Economically Disadvantaged—Combined Grades and Subjects



ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

LOCATIONS

Domestic

Washington, D.C.
Atlanta, GA
Baltimore, MD
Chapel Hill, NC
Chicago, IL
Columbus, OH
Concord, MA
Frederick, MD
Honolulu, HI
Naperville, IL
New York, NY
Portland, OR
Sacramento, CA
San Diego, CA
San Mateo, CA
Silver Spring, MD

INTERNATIONAL

Egypt
Ethiopia
Georgia
Haiti
Honduras
Kenya
Liberia
Malawi
Nicaragua
Pakistan
South Africa
Zambia

AIR®
AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000 | TTY: 877.334.3499

www.air.org

Making Research Relevant