Rivkin & Ishii Concerns

- We should worry about the validity of teacher ratings currently being used for rewards or sanctions
- Processes of students and teachers choosing schools and students and teachers being assigned to classrooms is not a random.

Decomposition of the teacher fixed effect

$$\hat{t}_j = t_j + \hat{\varepsilon}_i + \hat{\varepsilon}_p + \hat{\varepsilon}_s + \hat{\varepsilon}_n$$

- Teacher j fixed effect as a sum of the true teacher effect, the average student contribution to her own learning (subscript i), from her peer environment (subscript p), from the contribution of her school factors (subscript s), and random sampling error

Model of test score gains

$$(7)\quad A_{iGcsy} - A_{iG-1csy-1} = \gamma_i + \beta X_{iGy} + \tau P_{Gcsy} + \delta S_{Gcsy} + \lambda T_{Gcsy} + \eta_{iGcsy}$$

- student i in grade G in classroom c in school s in year y as a function of family background (X), peer composition in classroom c during year y (P), school factors specific to grade G in year y – including resources, principal quality, and school or district determined curriculum – (S), teacher quality (T), and a random error (e).

Achievement gain in grade G as a function of teacher quality (TQ) in grade G and a regression error.

$$\mathrm{A}_{iG} - A_{iG-1} = TQ_G \beta_{gain} + gain\ error = \tag{8}$$

$$TQ_G \beta_{gain} + \beta \sum_{g=1}^{G} (\theta^{g} - \theta^{g-1}) TQ_{G-g} + error$$

The expected value of the estimate of β_{gain} equals[7]

$$E(\hat{\beta}_{gain}) = \beta - \beta(\frac{(1-\theta^{G})\rho}{\mathrm{var}(TQ)} \tag{9}$$

Clever part

- Estimates whether systematic differences in the variance of teacher quality related to the ways in which
  - students and teachers are matched
    - Random vs. not random
  - structure of the value added model.
    - Gains vs. covariance adjustment

How do they measure sorting?

- Sample was dividing the sample into observations from schools/grade/year that exhibit significant signs of sorting and from those that do not.
  - Divides the sample based on whether observations from the school/grade/year reject the null hypothesis of no difference in the mean pretest score across classrooms.
  - Examines the transitions of students who remain in the school in grades g-1 and g and tests for the independence of the classroom allocation in the two grades using a chi square test.

Findings

- Estimates of the variance in teacher quality tend to be uniformly larger for schools in which the hypothesis of random sorting is rejected.
- Within school variance estimates from the test score gain models were much smaller than those from the lagged achievement model for the “non-random” sub-sample but not for the “random” sub-sample.

Rothstein (2007) falsification test

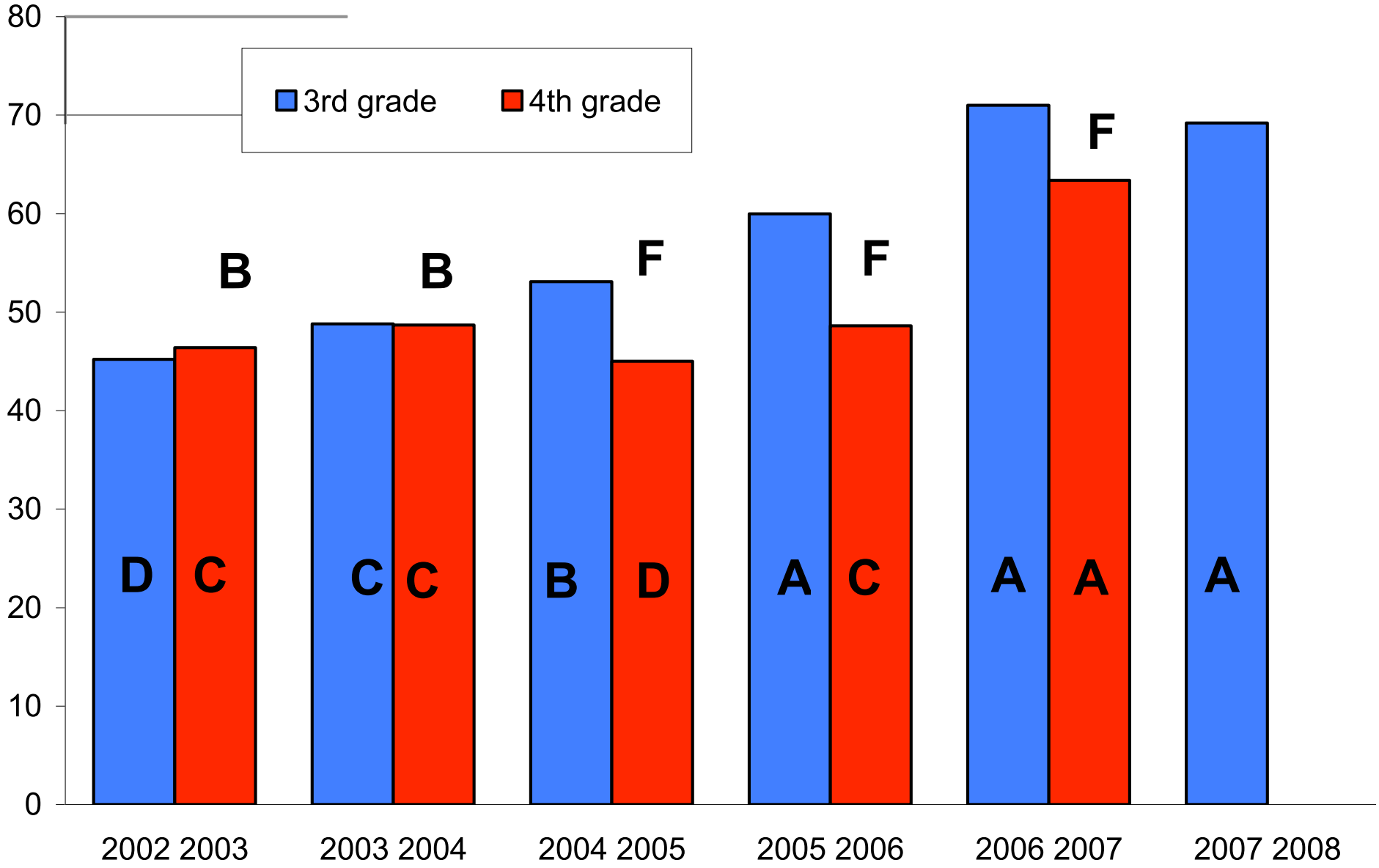
- Fixed effects for current year teacher are replaced by fixed for subsequent year teacher
  - The counterfactual within school variance estimates for future teachers using the lagged achievement model are roughly 60 percent smaller than the variance estimates using the actual teachers for grades in the non-random allocation category
  - The estimates of the counter-factual within school variances approach zero for the sample of schools in the “random” sample,

Why do Kane & Staiger and Rivkin & Ishii get different results?

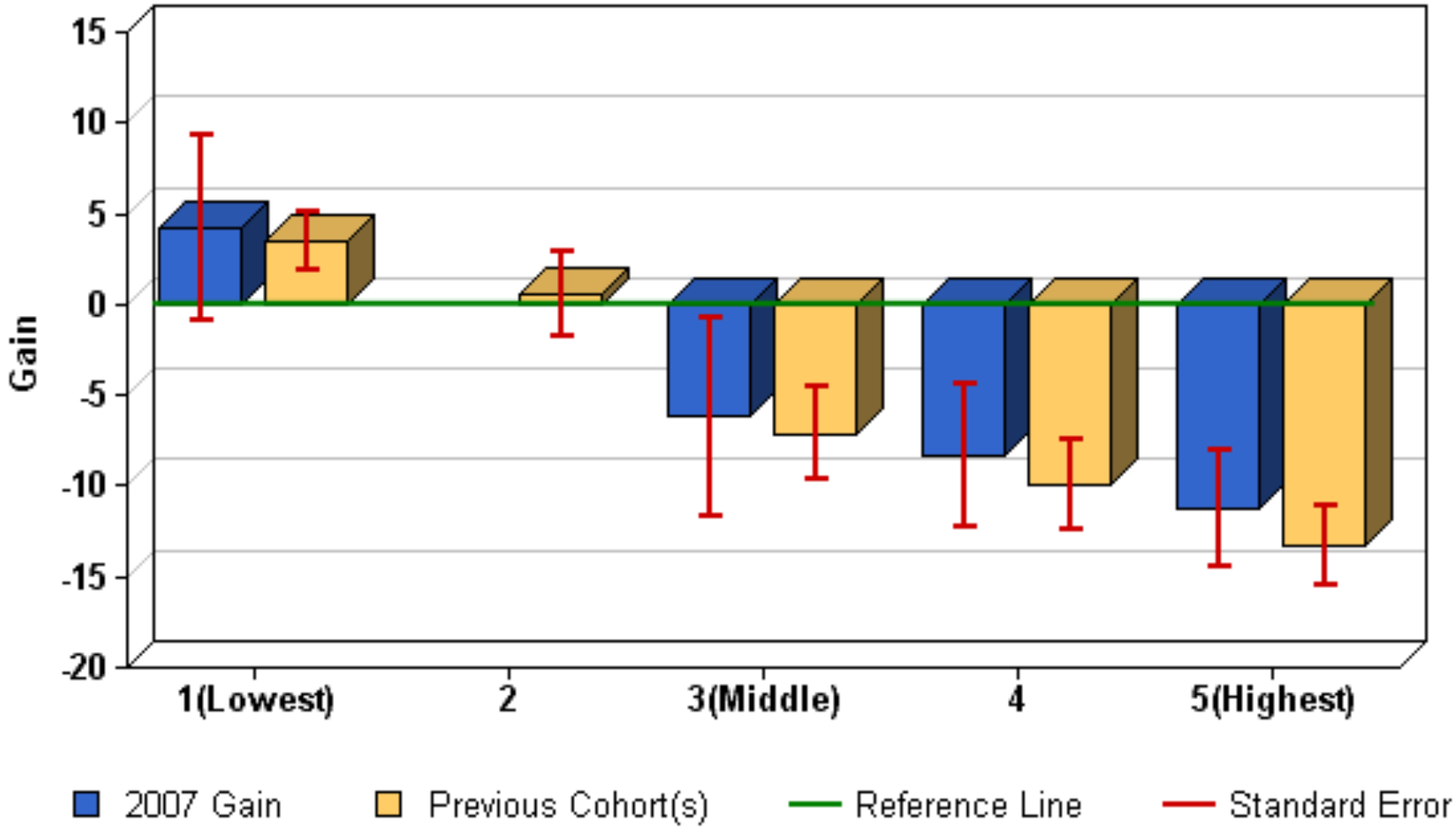
- More within school sorting of teachers to classes in middle school than elementary?
- Interactions between teacher skills and student ability/prior knowledge stronger in middle school than in elementary?

NCE Scores for 3rd and 4th graders in Mathematics

(50=State average score in that grade in 1998)

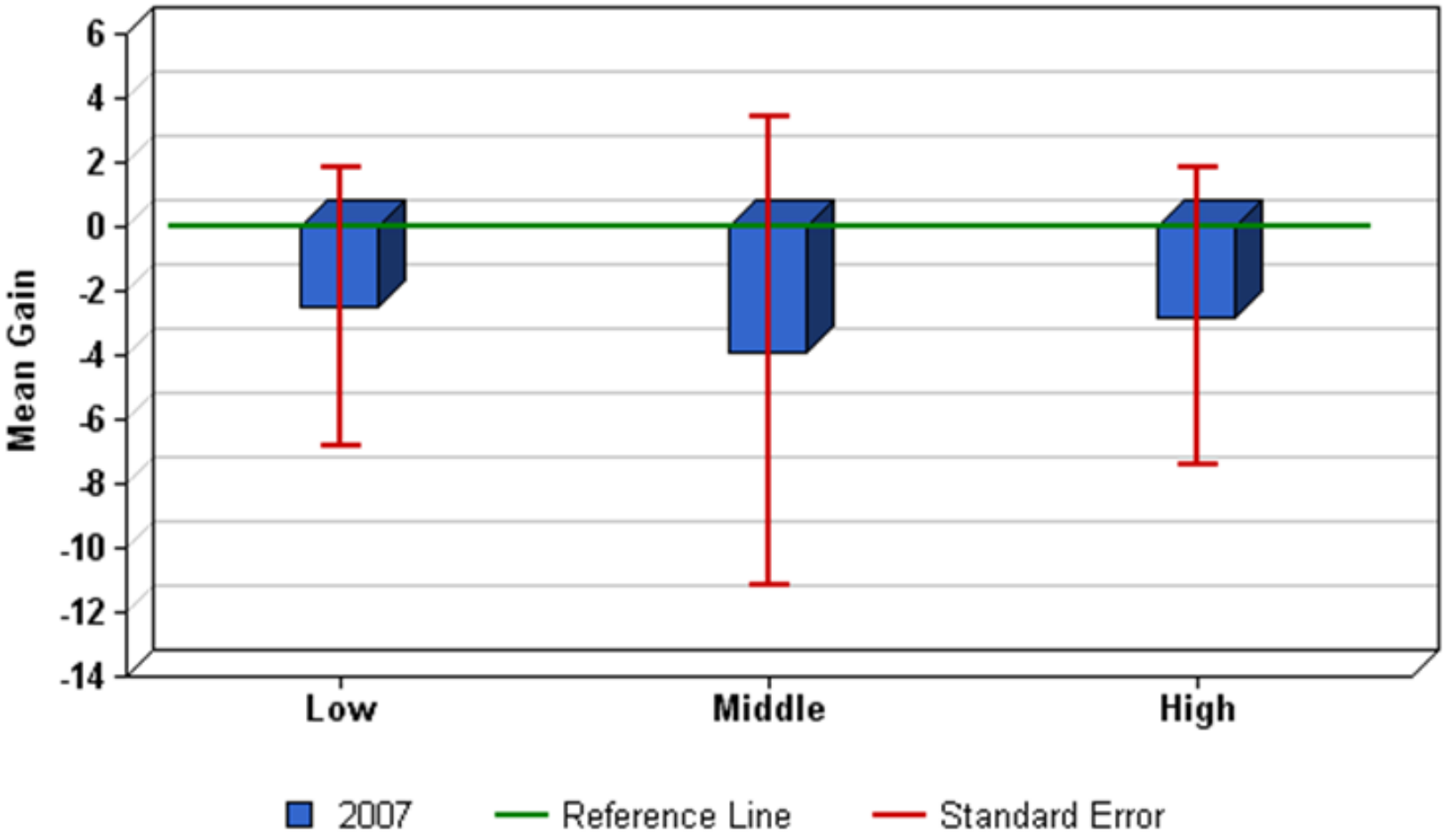


2007 Diagnostic Report for Elementary in Davidson County 4th Grade TCAP CRT Math



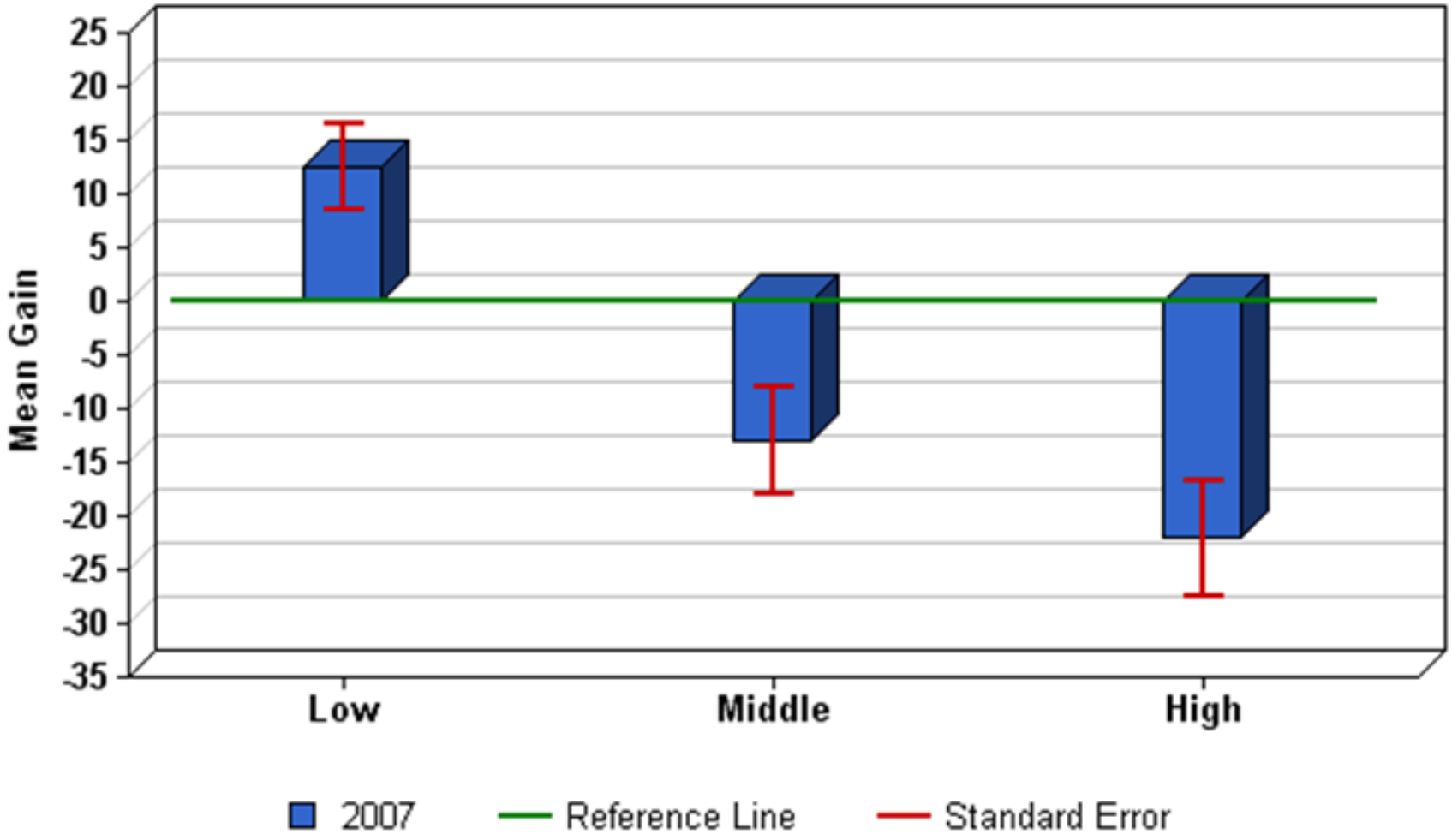
Class A

2007 Student Pattern Report for Elementary 4th Grade TCAP CRT Math



Class B

2007 Student Pattern Report for Elementary 4th Grade TCAP CRT Math



Class C

2007 Student Pattern Report for Elementary 4th Grade TCAP CRT Math