# How Does a Value-Added Model Compare to the Colorado Growth Model?

**October 10, 2013**

**Using Student Test Scores to Measure Teacher Performance: The State of the Art in Research and Practice**

**East Lansing, Michigan**

**Elias Walsh • Eric Isenberg**

# Colorado Growth Model (CGM)

- **Student growth percentiles (SGP)**
  - Ranks for students with similar prior achievement
  - Account for one to three same-subject pre-test scores
  - No demographic characteristics
- **Teacher evaluations**
  - CGM in eight states
  - Typically use median SGP for teacher

# Research Questions

- How large are differences between value-added and CGM scores?
- Are changes related to the characteristics of teachers' students?

# Literature Comparing CGM, Value Added

- **School-level evaluation scores**
  - High correlations, but substantive differences for some schools (Castellano 2011; Castellano and Ho 2012, 2013)
  - CGM benefits schools with fewer free and reduced-price lunch students (Ehlert et al. 2012)
- **Teacher-level evaluation scores**
  - CGM benefits teachers with fewer free and reduced-price lunch students (Wright 2010)
  - CGM also benefits teachers with higher-achieving students (Goldhaber et al. 2012)

# Our Contribution

- New context: high-stakes DC Public Schools (DCPS) IMPACT evaluation system
- How consequences for teachers change
  - Range from dismissal to bonus pay
  - Value added constituted 50 percent of evaluation
- New evidence on reasons for differences

# Value-Added Model

- Grades 4 through 8 in math and reading
- Accounts for same- and opposite-subject pre-test scores from prior year
- Accounts for demographic characteristics
- Pre-test measurement error correction (EIV)
- Teacher fixed effects
- Empirical Bayes shrinkage

# Data

- **2010–2011 DCPS students and teachers**
- **334 math teachers; 340 reading teachers**

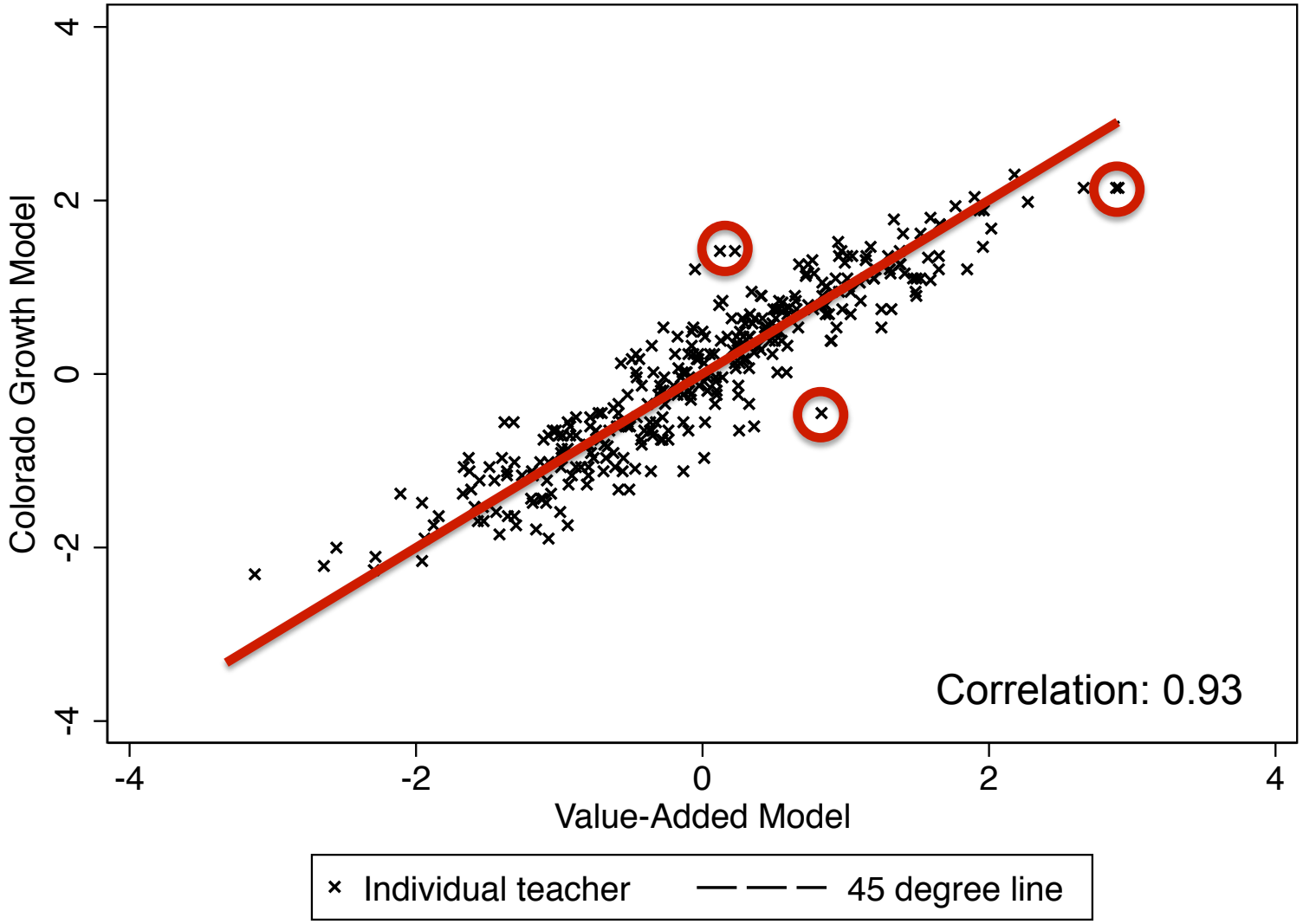
**10th and 90th percentiles of student disadvantage (teacher level)**

| | Low Level | High Level |
|---|---|---|
| Limited English proficiency | 0.0% | 33.0% |
| Learning disability | 0.0% | 20.9% |
| Eligible for free or reduced-price lunch | 23.4% | 96.1% |
| Pre-test score (standard deviations) | 0.73 | -0.65 |

# Why Evaluation Scores Might Change

- **No student demographic characteristics**
  - **May penalize teachers of disadvantaged students**
  - **Pre-test history could substitute**
- **No measurement error correction**
  - **May penalize teachers of low-achieving students**
  - **Correction attributes more achievement to pre-test**
- **No teacher fixed effects**
  - **May help teachers of disadvantaged students**
  - **Within- and between-teacher variation conflates**
    - Sorting of teachers based on student background
    - Relationship between background and achievement

# Comparison of Evaluation Scores for Math



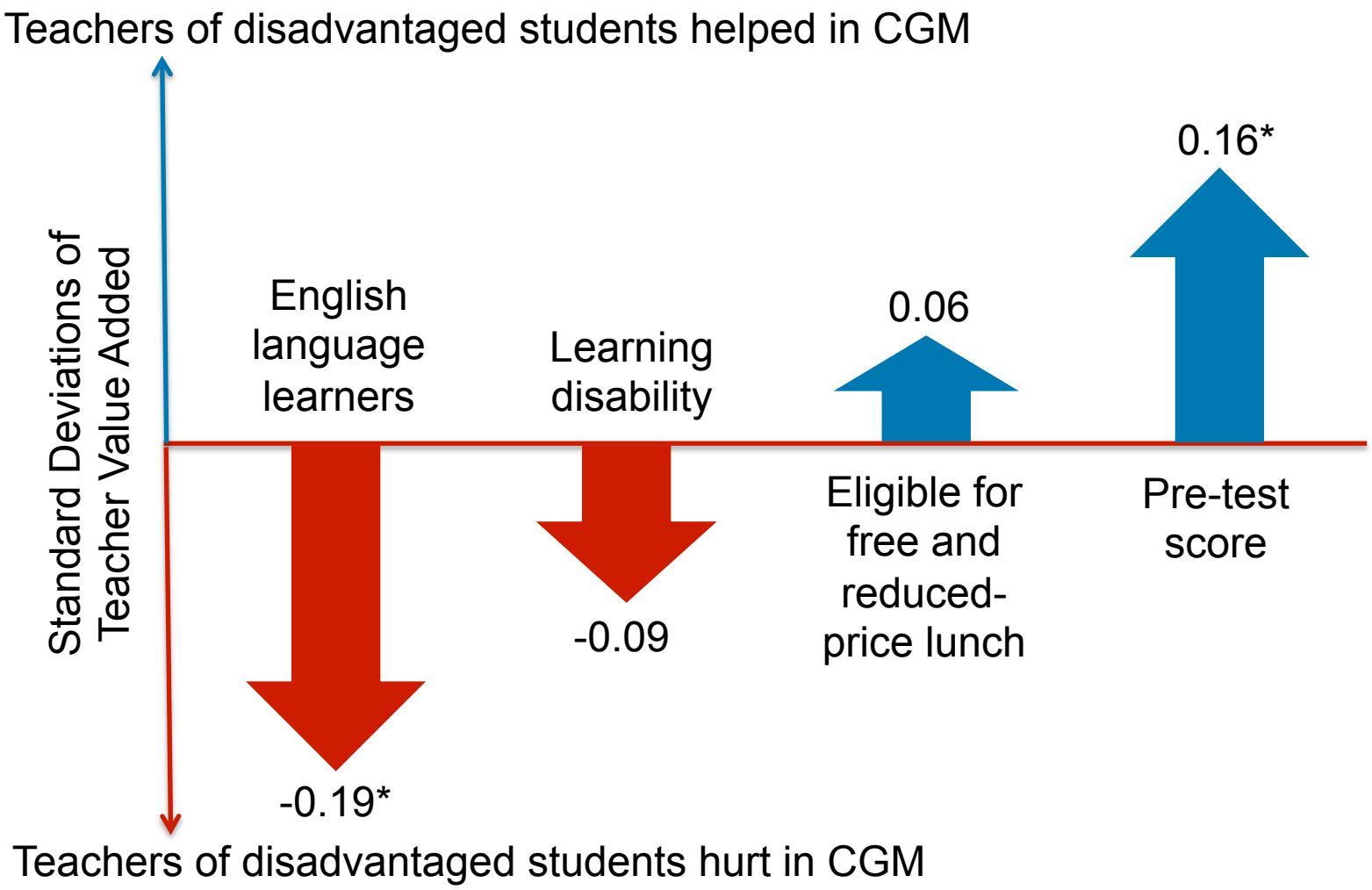



MATHEMATICA Policy Research

# Magnitude of Differences

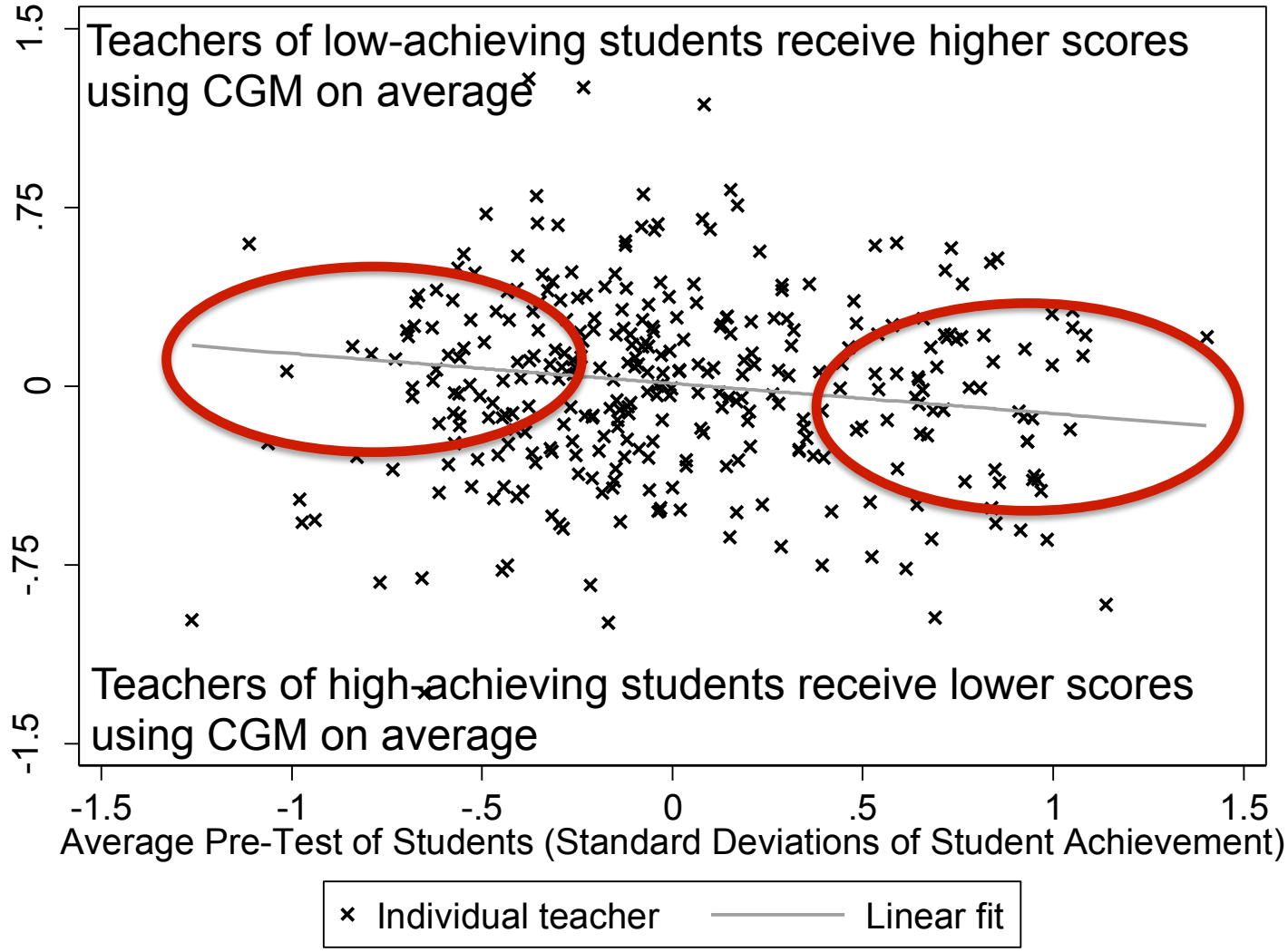
- **Median teacher moved**
  - 6 percentile ranks in math
  - 7 percentile ranks in reading
- **5 percent of teachers moved at least**
  - 22 percentile ranks in math
  - 25 percentile ranks in reading
- **14.2 percent changed IMPACT categories**

# Mixed Results: Teachers of Disadvantaged Students Helped or Hurt in CGM, Math




* Statistically significant at 5 percent level

# Change in Evaluation Scores by Average Pre-Test Score for Math



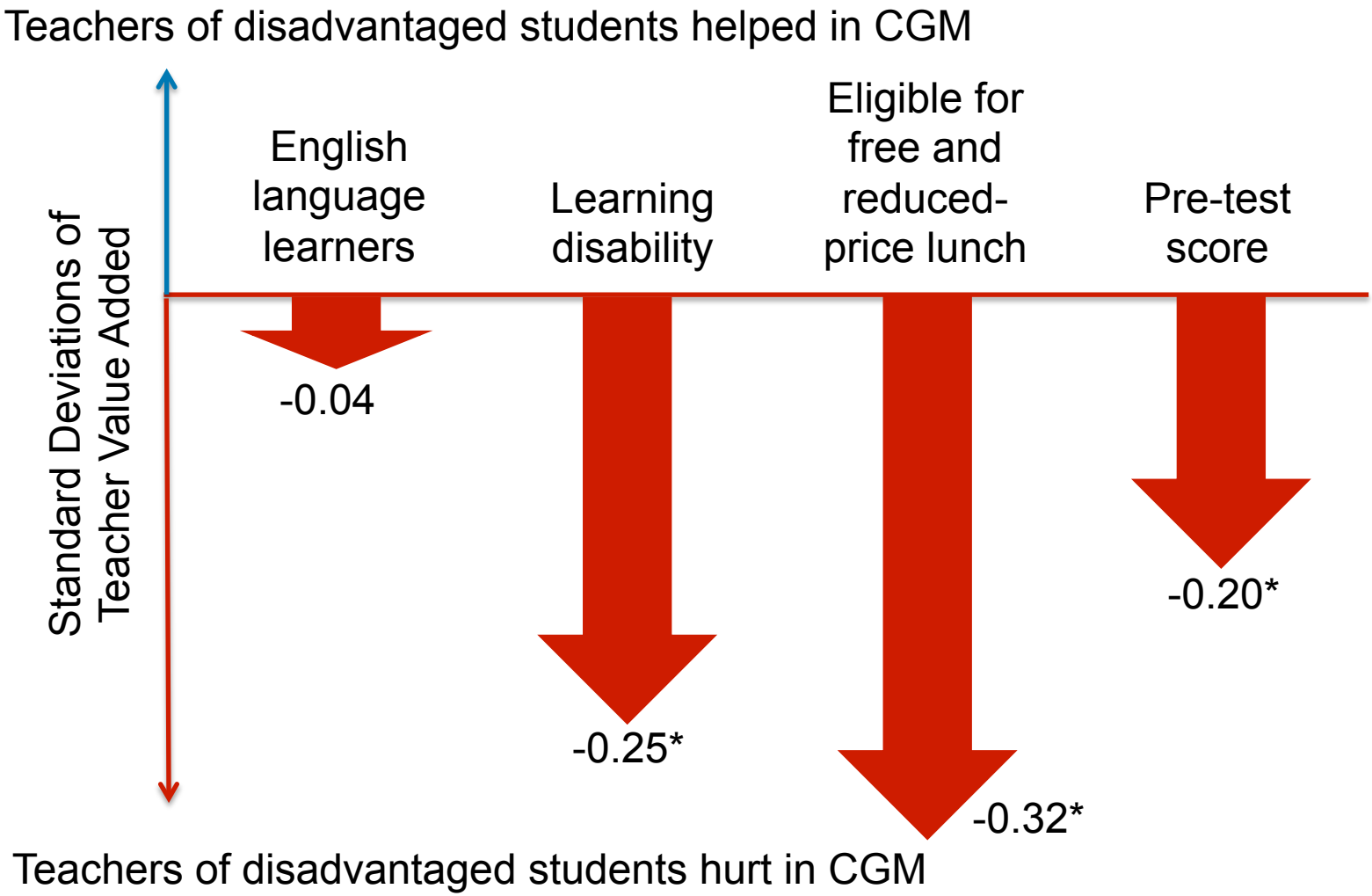



MATHEMATICA Policy Research

# CGM Helps Teachers of Low-Achieving Students

- **No fixed teacher effects can explain if**
  - **Less effective teachers matched to more disadvantaged students**
  - **Effect is larger than other consequences of CGM**
- **Finding contrasts with previous work**
- **Comparing CGM with average-residuals model flips results**

# Teachers of Disadvantaged Students in CGM Versus Average-Residuals Model, Math




* Statistically significant at 5 percent level

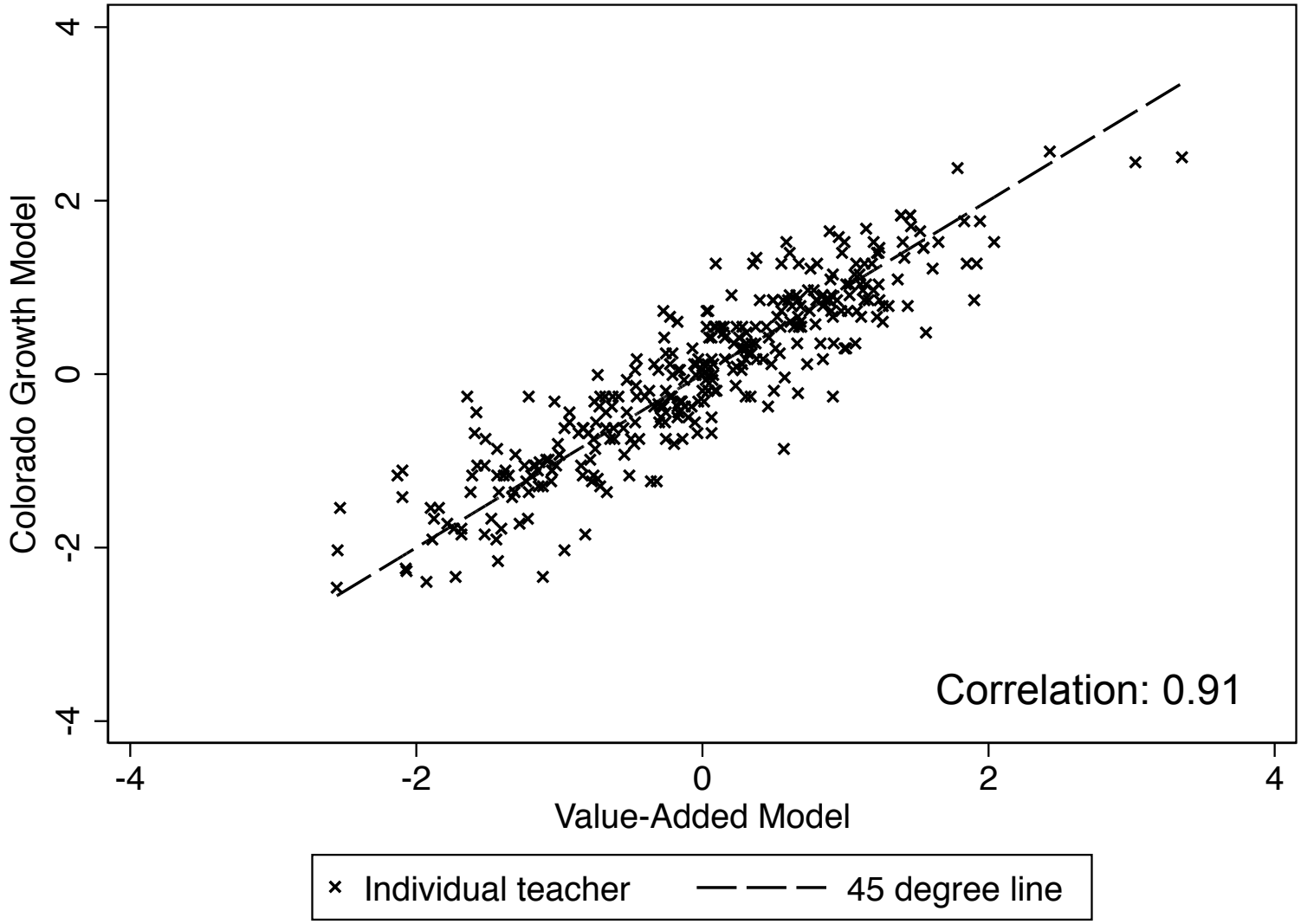
# Summary and Conclusion

- 14 percent of teachers would have different consequences
- Mixed results for teachers of disadvantaged students
  - Lower pre-test scores: tend to do better with CGM
  - More English language learners: tend to do worse
  - Most changes not related to student characteristics
- Changes raise concerns about bias in CGM

# For More Information

- **Please contact:**
  - **Elias Walsh**
    - ewalsh@mathematica-mpr.com
  - **Eric Isenberg**
    - ejisenberg@mathematica-mpr.com

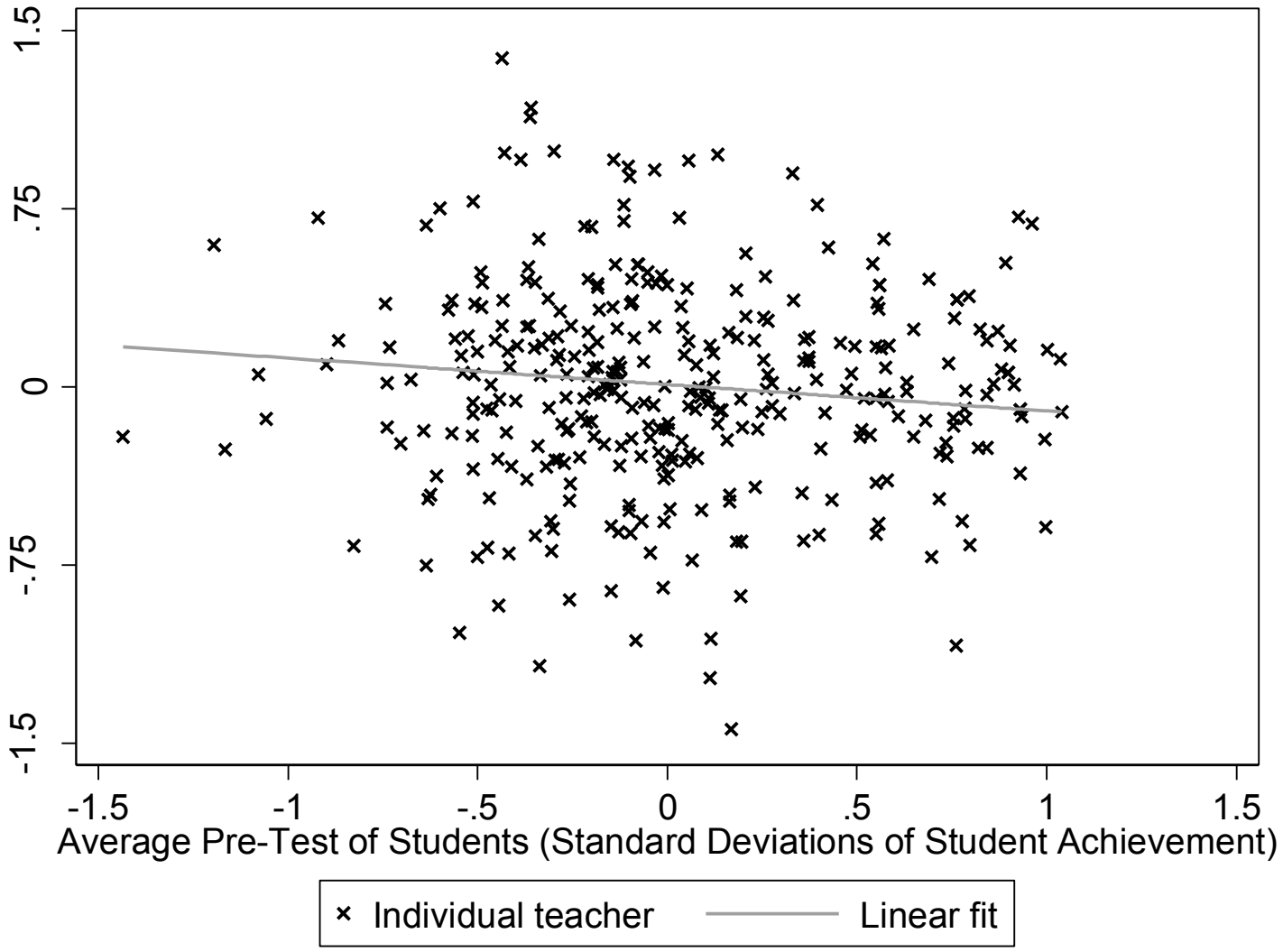
# Comparison of Evaluation Scores for Reading



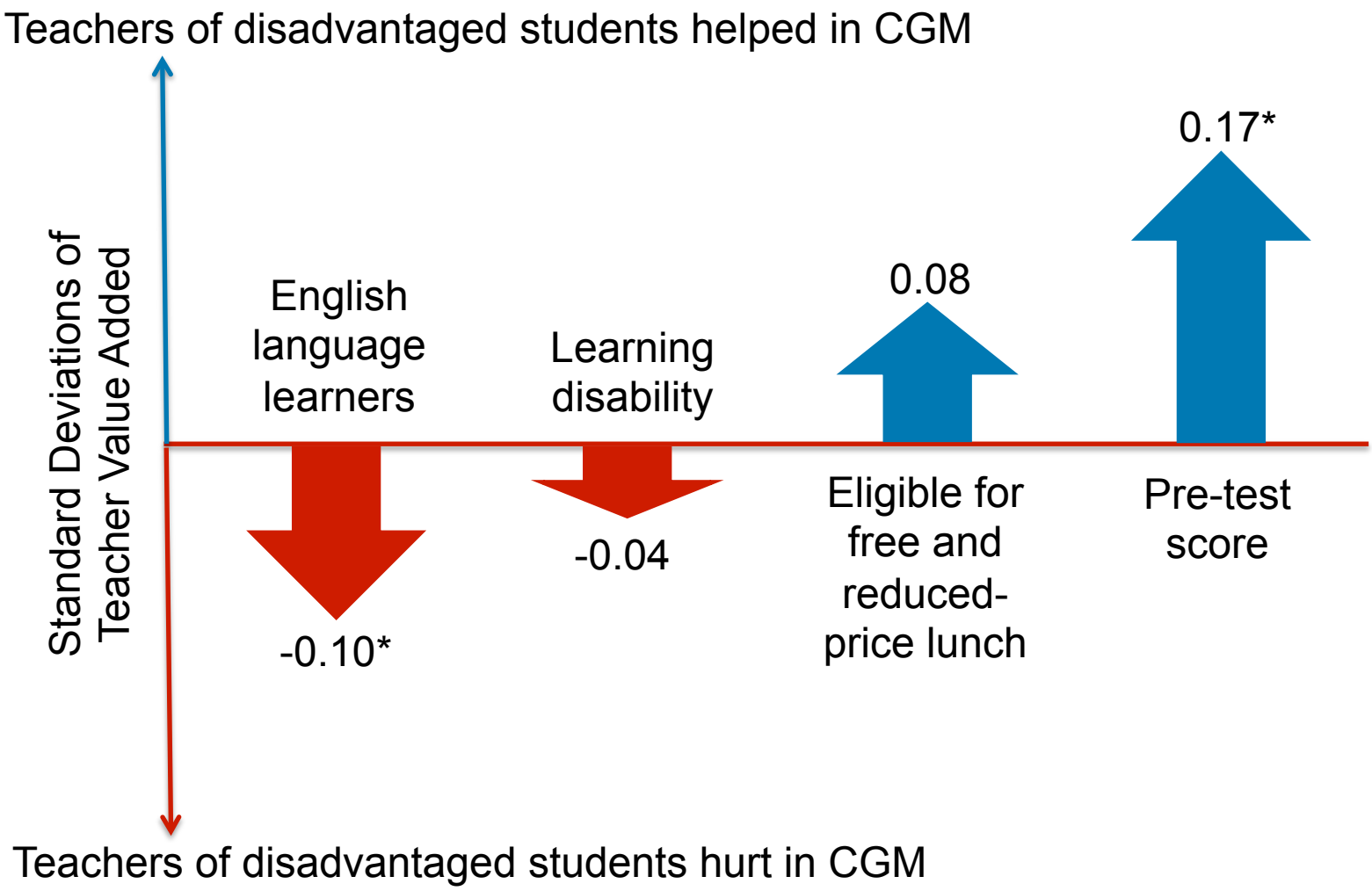



MATHEMATICA Policy Research

# Change in Evaluation Scores by Average Pre-Test Score for Reading

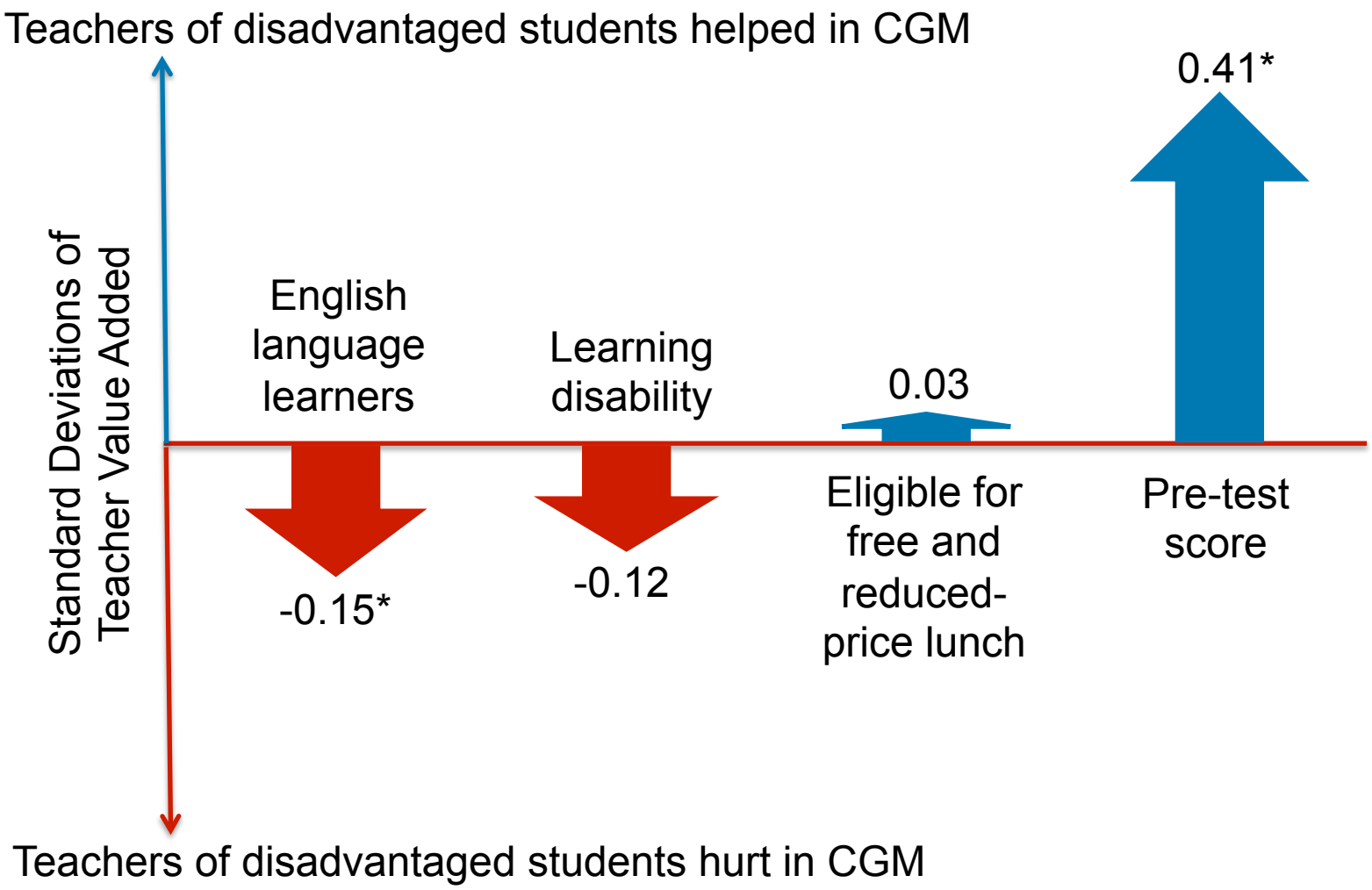


# Mixed Results: Teachers of Disadvantaged Students Helped or Hurt in CGM, Reading



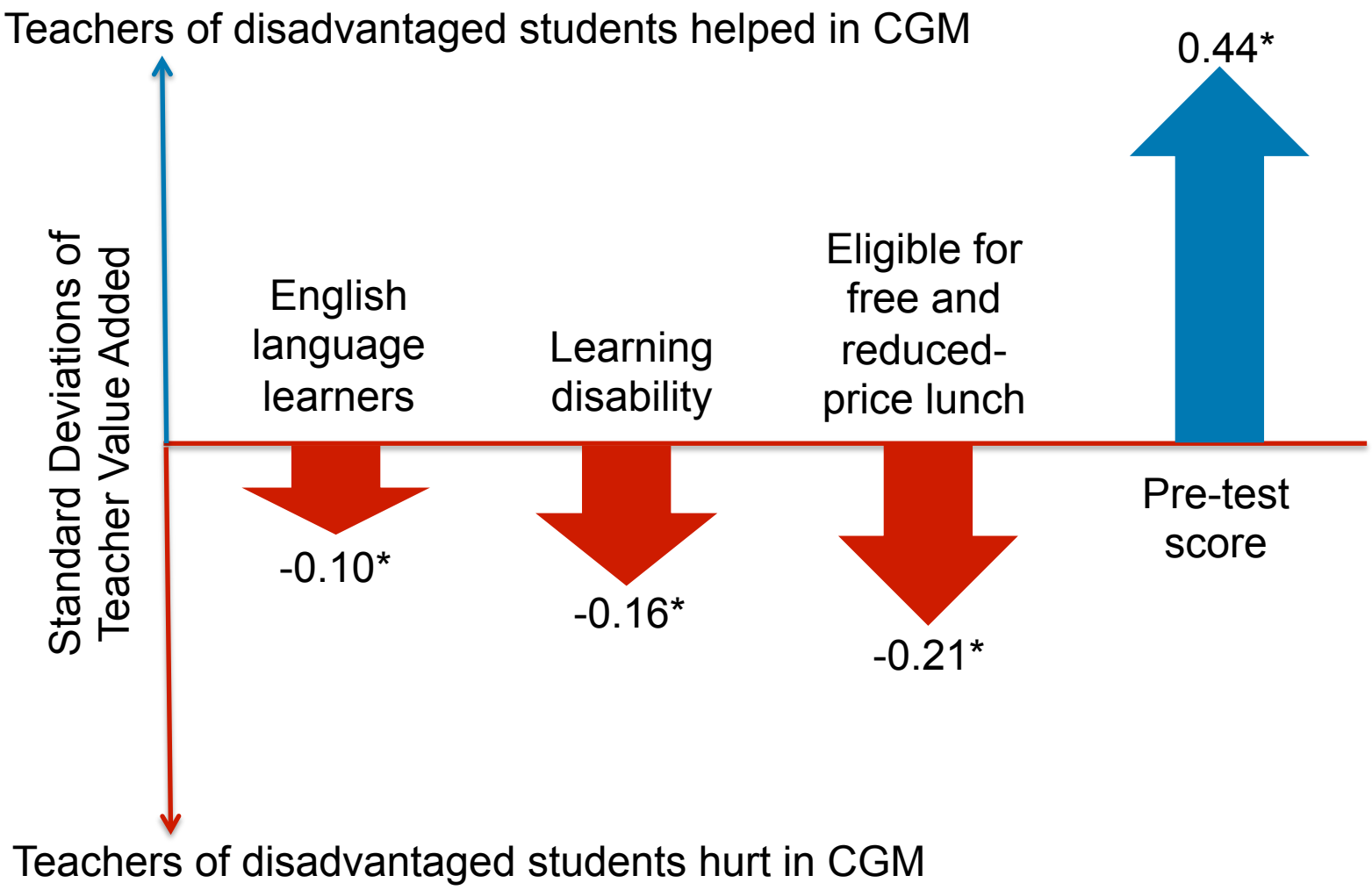

* Statistically significant at 5 percent level

# Marginal Differences for Teachers of Disadvantaged Students in CGM, Math




* Statistically significant at 5 percent level

# Marginal Differences for Teachers of Disadvantaged Students in CGM, Reading




* Statistically significant at 5 percent level