NATIONAL CENTER for ANALYSIS of LONGITUDINAL DATA in EDUCATION RESEARCH
RESEARCH FOR POLICY

Teacher Quality Gaps and Student Outcomes: Assessing the Association Between Teacher Assignments and Student Math Test Scores and High School Course-Taking

Dan Goldhaber
American Institutes for Research
University of Washington

Roddy Theobald
American Institutes for Research

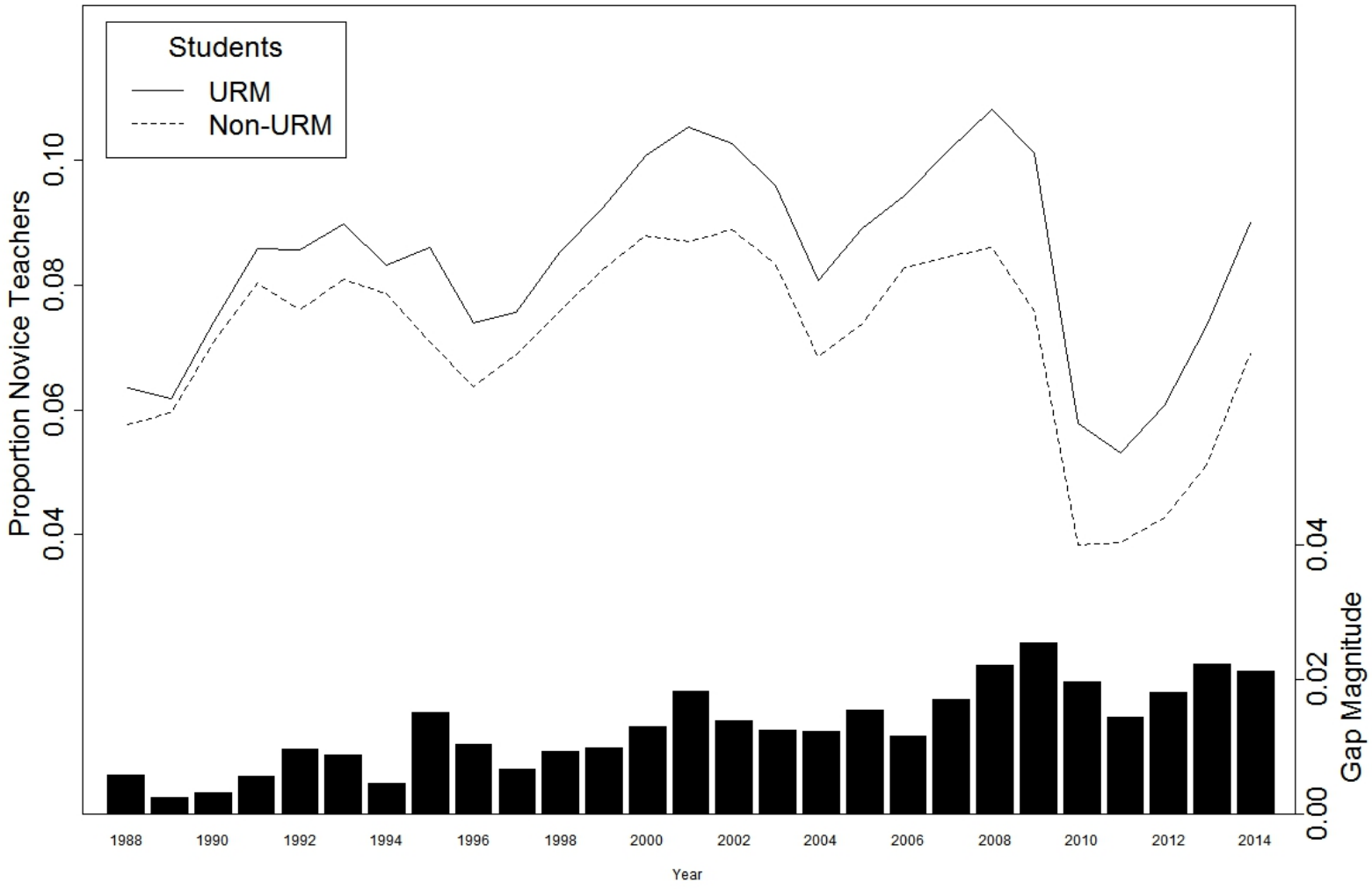
Danielle Fumia
Washington State Institute for Public Policy

AIR® American Institutes for Research®

Background

- Considerable prior evidence documents persistent gaps in test achievement, high school advanced course taking, & educational attainment
- Teacher quality has large impact on student test and non-test outcomes
  - The implications of assignment to very effective vs. ineffective teachers is estimated to be equivalent to more than a year's worth of learning (Hanushek, 1992) and "replacing a teacher whose VA is in the bottom 5% with an average teacher would increase the present value of students' lifetime income by approximately $250,000 per classroom."(Chetty et al., 2014)
- Significant evidence of "teacher quality gaps" (TQGs) between advantaged and disadvantaged students in U.S. public schools in terms of *teacher credentials* (e.g. Clotfelter et al., 2005; Kalogrides and Loeb, 2013; Lankford et al., 2002)
  - More mixed *characterizations* of evidence of TQGs in terms of *value added* (e.g., Goldhaber et al., 2015; Isenberg et al., 2013, 2016**;** Mansfield, 2015; Sass et al., 2010; Steele et al., 2015)
- Large impacts of teachers, combined with evidence of a considerable amount of heterogeneity in teacher effectiveness, arguably make teachers both the key *schooling* variable influencing student outcomes and central to the equality of educational opportunity

Last Year's CALDER Conference



What Do We Do Here?

- We document the extent to which teacher assignments in grades 4-8 are *associated* with 8th grade math test scores and high school course-taking
  - Note the word *associated*
- We address three questions:
  1. To what extent are there persistent TQGs between advantaged and disadvantaged students in grades 4-8?
     - Define advantaged and disadvantaged by underrepresented minority (URM; African-American, American Indian/Alaskan Native, & Hispanic) and free/reduced price lunch (FRL) status
  2. To what extent does teacher 4th – 8th grade assignment (indicators or VA) explain end-of-grade 8th grade math tests and advanced HS math courses?
  3. By how much to advantaged-disadvantage outcome gaps shrink when accounting for teacher assignment?

Quick Questions and Answers

1. To what extent are there persistent TQGs between advantaged and disadvantaged students in grades 4-8?
   - Consistent evidence of persistent TQGs in grades 4 & 5; grades 6-8 answer is sensitive to value-added specification
2. To what extent does teacher 4th – 8th grade assignment (indicators or VA) explain end-of-grade 8th grade math tests and advanced HS math courses?
   - Teacher assignment is a highly significant predictor of 8th grade test score and HS math course taking
   - 4th-8th grade value-added statistically significant
3. By how much to advantaged-disadvantaged outcome gaps shrink when accounting for teacher assignment?
   - Direct controls for teacher assignments shrink 8th grade test gap for URM and FRL by 16% and 21% respectively; shrink course-taking gap by 33% for both
   - Accounting for estimates from preferred VA specification shrinks 8th grade test gap by 8-9% for both; accounting for VA shrinks course-taking gap by about 7% for both

Data

- Provided by the Washington State Office of Superintendent of Public Instruction (OSPI)
  - 11 years of administrative data (2005-06 to 2015-16)

- Analytic sample:
  - 8th Grade Test Sample: six cohorts of students: 330,539 student observations and 36,739 teacher-year observations (11,194 unique teachers)
  - High School Course-Taking Sample: two cohorts of students: 104,001 student observations and 12,829 teacher-year observation (7,874 unique teachers)

Implications of Sample Restriction

	Unrestricted sample	8th Grade Achievement Analytic Sample					High School Course Taking Analytic Sample				
		All	URM	non-URM	FRL	non-FRL	All	URM	non-URM	FRL	non-FRL
	1	2	3	4	5	6	7	8	9	10	11
Baseline characteristics											
URM student	0.258	0.256	1.000	0.000	0.447	0.106	0.225	1.000	0.000	0.443	0.092
3rd Grade FRL	0.449	0.441	0.769	0.328	1.000	0.000	0.379	0.746	0.273	1.000	0.000
3rd Grade Math Score	0.000	0.032	-0.421	0.188	-0.316	0.307	0.095	-0.402	0.240	-0.289	0.330
	(1.000)	(0.977)	(0.934)	(0.942)	(0.931)	(0.923)	(0.963)	(0.962)	(0.914)	(0.951)	(0.892)
3rd Grade Reading Score	0.000	0.026	-0.408	0.176	-0.328	0.306	0.084	-0.385	0.221	-0.302	0.320
	(1.000)	(0.976)	(0.936)	(0.945)	(0.940)	(0.912)	(0.970)	(0.961)	(0.929)	(0.960)	(0.898)
Female	0.490	0.492	0.495	0.491	0.495	0.490	0.499	0.507	0.496	0.507	0.493
3rd Grade Special Education	0.127	0.118	0.124	0.116	0.141	0.100	0.105	0.113	0.102	0.126	0.092
3rd Grade English Language Learner	0.094	0.096	0.278	0.033	0.190	0.021	0.091	0.297	0.030	0.207	0.019
3rd Grade Gifted	0.030	0.029	0.009	0.036	0.009	0.045	0.031	0.011	0.037	0.010	0.044
N	437123	330539	84744	245795	145811	184728	104001	23438	80563	39464	64537

Unrestricted Restricted Sample Differential = -0.032

URM non-URM Differential = -0.609

FRL non-FRL Differential = -0.623

Teacher Experience Gaps

	Unrestricted sample	8th Grade Achievement Analytic Sample					High School Course Taking Analytic Sample				
		All	URM	non-URM	FRL	non-FRL	All	URM	non-URM	FRL	non-FRL
	1	2	3	4	5	6	7	8	9	10	11
Teacher experience											
4th Grade <= 2 Years Experience	0.048	0.051	0.065	0.047	0.056	0.048	0.068	0.082	0.064	0.075	0.063
5th Grade <= 2 Years Experience	0.038	0.040	0.050	0.037	0.045	0.036	0.048	0.057	0.046	0.054	0.044
6th Grade <= 2 Years Experience	0.034	0.037	0.043	0.034	0.038	0.035	0.024	0.026	0.023	0.024	0.024
7th Grade <= 2 Years Experience	0.041	0.044	0.055	0.040	0.049	0.039	0.046	0.061	0.042	0.055	0.041
8th Grade <= 2 Years Experience	0.038	0.044	0.052	0.042	0.047	0.042	0.038	0.048	0.035	0.043	0.035

- Probability of disadvantaged students having inexperienced teachers greater in every grade
- Dosage effects: URM students are 50% more likely than non-URM students to have two or more novice teachers in 4^{th}-8^{th} grades

Estimating Value Added

- Value added, what are we worried about:
  1. No circularity in estimates of association/impact of teachers
     - No problem: leave-one-out method
  2. Properly account for “drift” of value added
     - No problem: estimate forecasting coefficients
  3. Properly account for value of the experience of teachers when students have them:
     - No problem: estimate returns to experience and add experience back in
  4. Disentangle peer effects from differences in teacher quality across classrooms
     - Ah, this one isn’t so easy or clear
       - Control for just individual covariates or also classroom covariates in VAM? One accounts for potential of peer effects, but may attribute true differences in teacher quality to classroom covariates

Measured Value Added Gaps

Table 2. Correlations between VAM estimates with and without classroom controls

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Correlation	0.992	0.983	0.971	0.883	0.914

Measured Value Added Gaps

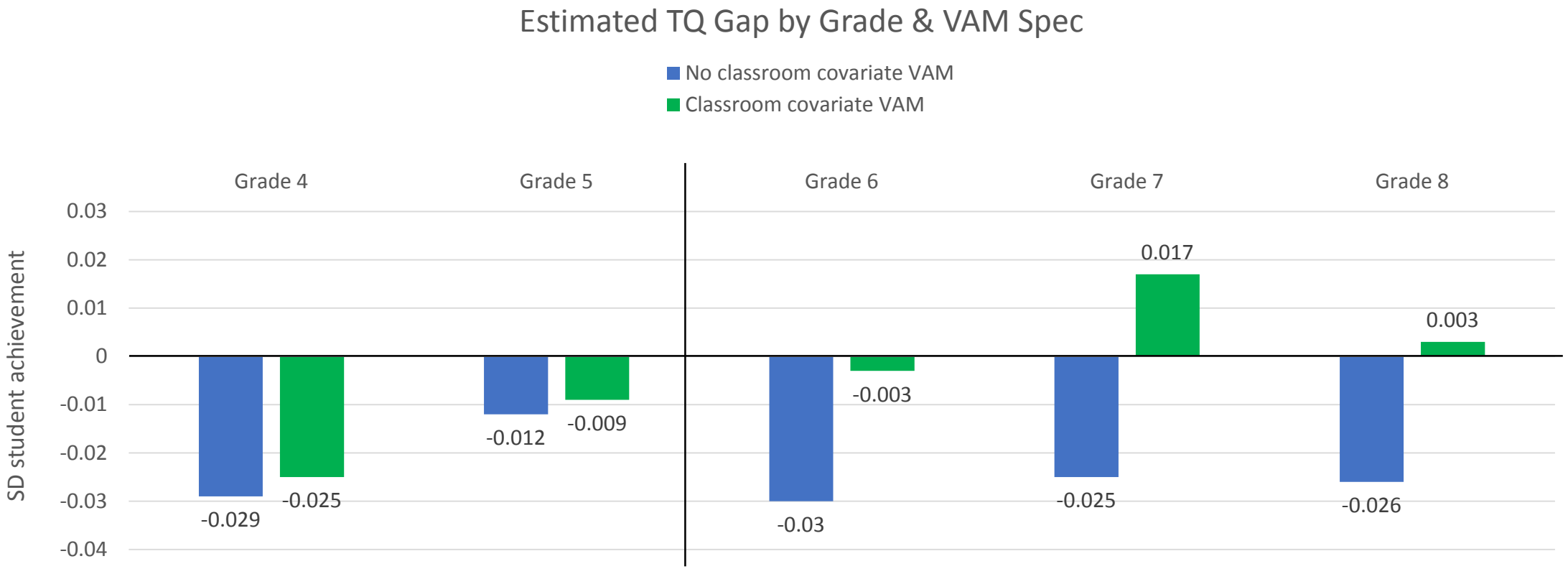
Table 2. Correlations between VAM estimates with and without classroom controls

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Correlation	0.992	0.983	0.971	0.883	0.914

Measured Value Added Gaps

Table 2. Correlations between VAM estimates with and without classroom controls

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Correlation	0.992	0.983	0.971	0.883	0.914



Differences in Outcomes

	Unrestricted sample	8th Grade Achievement Analytic Sample					High School Course Taking Analytic Sample				
		All	URM	non-URM	FRL	non-FRL	All	URM	non-URM	FRL	non-FRL
	1	2	3	4	5	6	7	8	9	10	11
Outcomes											
8rd Grade Math Score	0.042	0.066	-0.374	0.218	-0.301	0.356	0.155	-0.279	0.286	-0.200	0.378
	(0.989)	(0.978)	(0.880)	(0.965)	(0.896)	(0.943)	(0.959)	(0.858)	(0.950)	(0.882)	(0.939)
N	344537	330539	84744	245795	145811	184728	97635	22605	75030	37667	59968
Any advanced math course in HS	0.419	0.408	0.292	0.443	0.279	0.489	0.421	0.296	0.457	0.285	0.504
Number of advanced math courses in high school	0.631	0.597	0.393	0.659	0.372	0.739	0.634	0.404	0.702	0.385	0.787
	(0.877)	(0.843)	(0.692)	(0.874)	(0.672)	(0.906)	(0.878)	(0.706)	(0.912)	(0.690)	(0.944)
N	106525	97635	22605	75030	37667	59968	104001	23438	80563	39464	64537
8rd Grade Math Score	0.042	0.066	-0.374	0.218	-0.301	0.356	0.155	-0.279	0.286	-0.200	0.378

URM non-URM Differential = -0.592

FRL non-FRL Differential = -0.657

URM non-URM Differential = -0.298

FRL non-FRL Differential = -0.402

Dosage Models: Analytic Approach

- Estimate models that predict 8th-grade math test scores or number of advanced high school math courses as a function of baseline 3rd-grade covariates (e.g., test scores and demographics) and:
  1. Nothing about teacher assignments
  2. Individual teacher *indicators* from 4th-8th grade
  3. Individual teacher *value added* from 4th-8th grade
- Differences between 1 and 2 give importance of teacher assignments
- Differences between 1 and 3 give importance of measured teacher value added

Summary of Findings

- Estimates of TQGs are sensitive to VA specification
  - Disagreement at the middle school level, despite correlations of two specifications near 0.9
- Huge drop in URM and FRL coefficients associated w/ addition of 3rd grade test scores
  - Roughly 2/3 reduction in estimated URM and FRL gaps
  - Covariate adjustments show URM students taking more advanced courses
- Teacher assignment (F-tests) and value added are highly predictive of both 8th grade achievement and high-school course-taking
  - Even teachers in early grades matter
- Classroom covariate and non-classroom covariate VAMs disagree
  - Classroom covariate VA shows that equality of teacher VA assignment would lead to *larger* student outcome gaps – counterintuitive, but consistent with estimates of TQGs

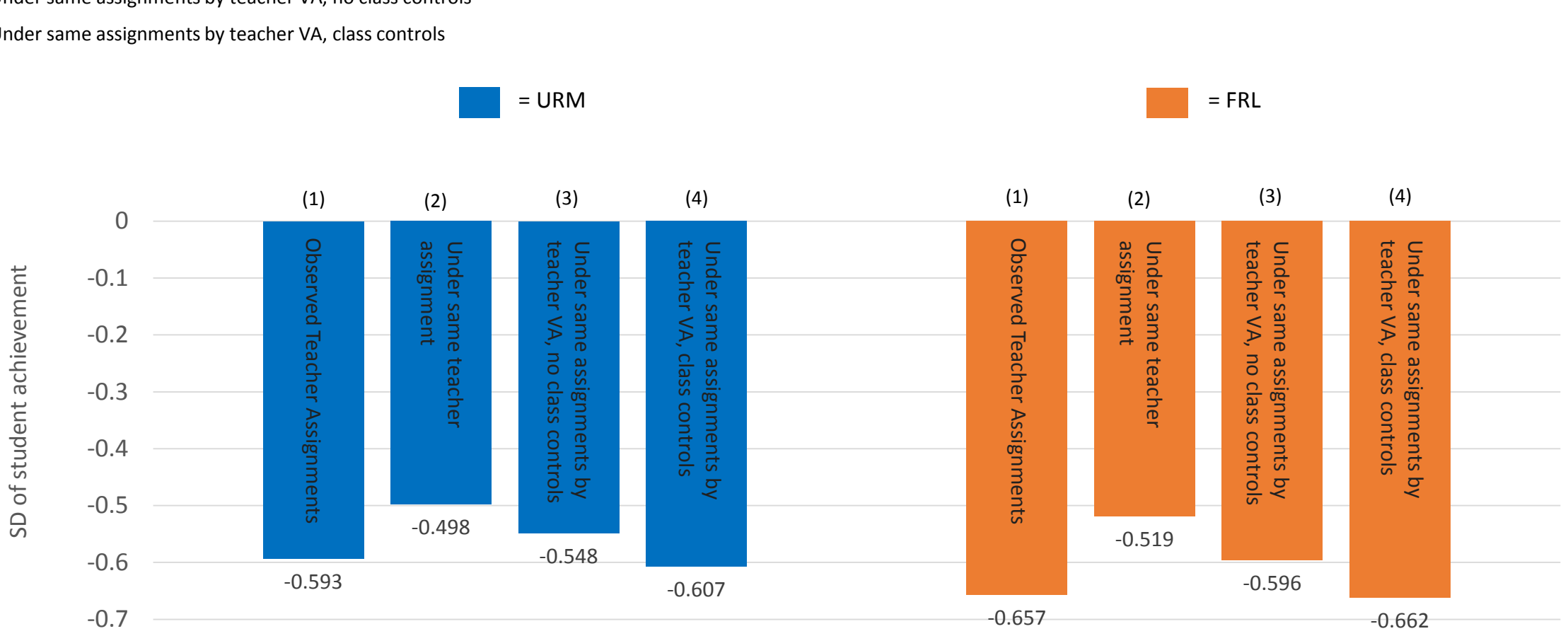
High Predicted TQ Effects on Student Outcome Gaps

(1) Observed teacher assignments

(2) Under same teacher assignment

(3) Under same assignments by teacher VA, no class controls

(4) Under same assignments by teacher VA, class controls



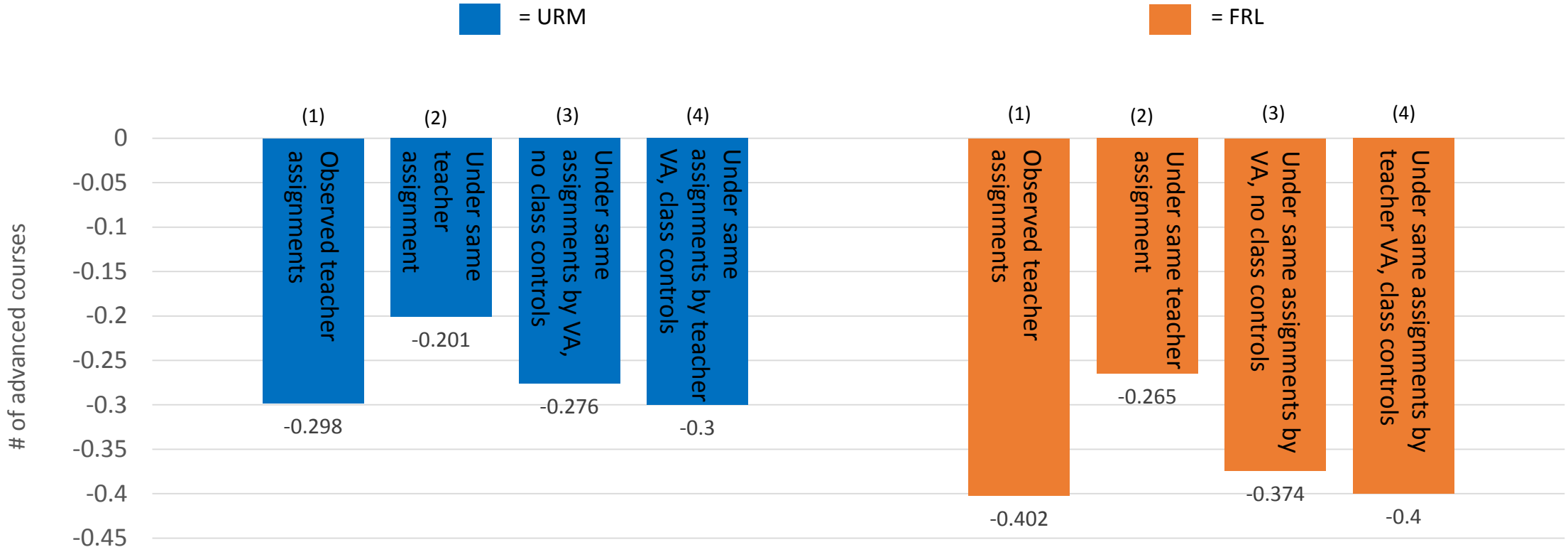
High Predicted TQ Effects on Student Outcome Gaps

(1) Observed teacher assignments

(2) Under same teacher assignment

(3) Under same assignments by teacher VA, no class controls

(4) Under same assignments by teacher VA, class controls



Conclusions

- Much of the gaps we observe in 8th grade tests and high school course-taking can be explained by achievement differentials that existed in the 3rd grade
  - Policymakers wishing to alleviate later achievement gaps either need to intervene earlier in a student's academic career or be far more aggressive after the 3rd grade to address academic deficiencies
- Findings support the importance of considering teacher assignment as one policy lever for addressing achievement gaps
  - Also support the importance of value added, but show the need for more research on value-added specification in middle schools
- But while teacher quality gaps appear to be an important contributor to gaps in longer-term student outcomes, eliminating them should not be viewed as a panacea for eliminating student outcome gaps

8th Grade Test Score Models

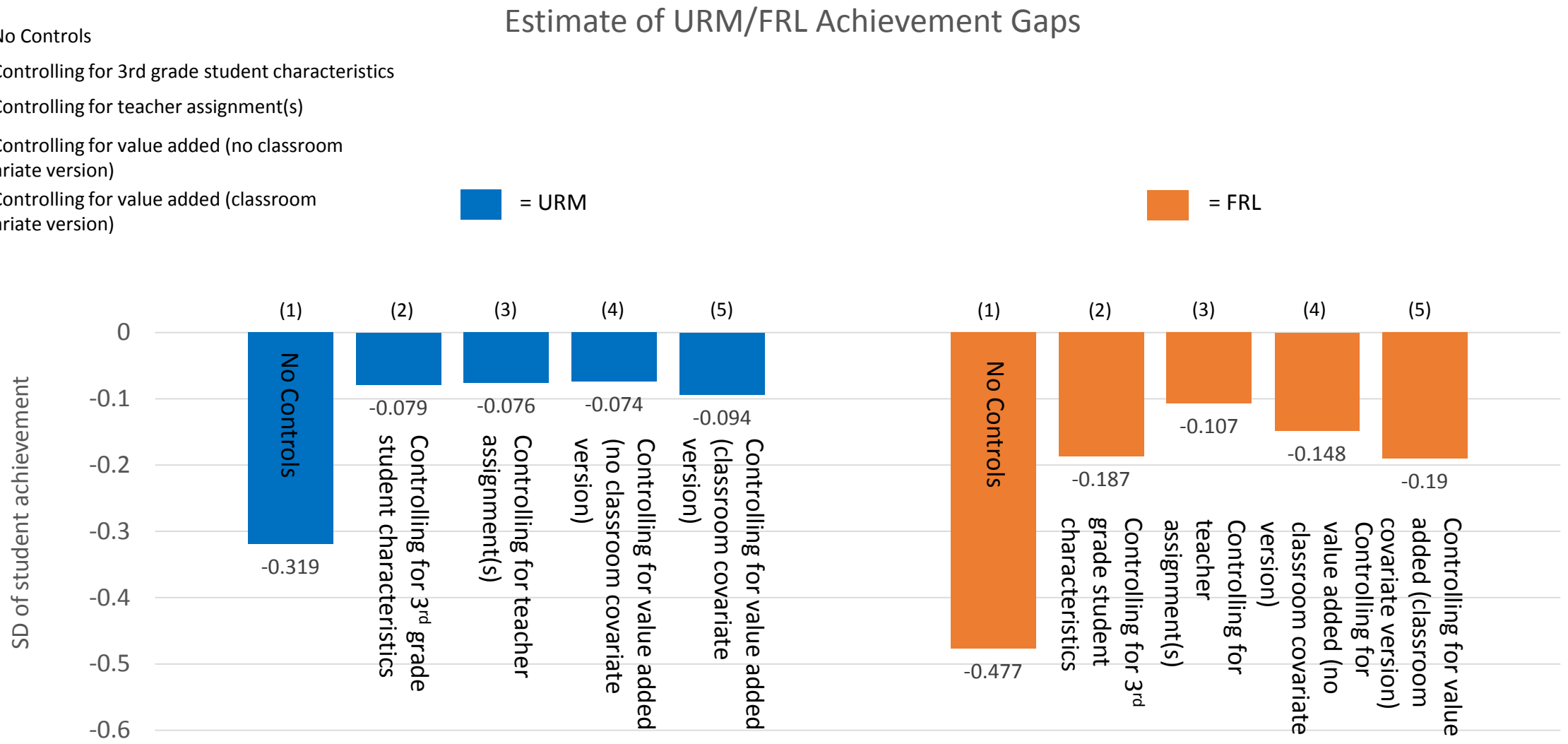
(1) No Controls

(2) Controlling for 3rd grade student characteristics

(3) Controlling for teacher assignment(s)

(4) Controlling for value added (no classroom covariate version)

(5) Controlling for value added (classroom covariate version)



High School Course-Taking Models

(1) No Controls

(2) Controlling for 3rd grade student characteristics

(3) Controlling for teacher assignment(s)

(4) Controlling for value added (no classroom covariate version)

(5) Controlling for value added (classroom covariate version)