The Long-Term Impacts of Teachers:
Teacher Value-Added and Students' Outcomes in Adulthood

Raj Chetty, Harvard
John N. Friedman, Harvard
Jonah Rockoff, Columbia

October 2012

Introduction: Teacher Value-Added

- What is the best way to measure and improve teacher quality?
- One approach: "value-added" (VA) measures [Hanushek 1971, Murnane 1975,...]
  - Rate teachers based on their students' test score gains
- School districts have started to use such VA measures, leading to considerable debate in policy circles
- Debate about VA stems largely from three key issues [Kane and Staiger 2008, Rothstein 2010, Darling-Hammond et al. 2012]

Question 1: Are VA Measures Accurate?

- Teachers are assigned different types of students

  → Teachers' estimated VA may depend more on the types of students they get rather than the teachers' actual impacts

- Standard approach attempts to account for this problem by adjusting for student characteristics
  - But is this sufficient to obtain accurate estimates of teacher impacts?

- Resolving this issue is critical for policy [Rothstein 2010]
  - Does VA unfairly penalize teachers for their mix of students?

Question 2: Does VA Predict Long-Term Impacts?

- Even if teacher VA is an accurate measure of teachers' impacts on test scores, it may not be a good proxy for teacher quality
  - Do high VA teachers improve students' long-term success?

    ... or are they simply better at teaching to the test?

Question 3: Is VA too Unstable for Evaluation?

- Teacher VA estimates fluctuate across years because they are based on samples with relatively few students
  - Many other influences on student progress → noise in VA estimates
  - Are VA estimates based on a few years of data sufficiently stable to be useful for teacher evaluation?

Data

- We answer these two questions by tracking one million children from childhood to early adulthood

- Implement analysis by linking two large databases
  - Data from a large school district: teachers, class assignments, and test scores from 1991-2009 for students in grades 3-8
  - Administrative tax records on student outcomes in adulthood (earnings, college, teenage birth) and parent characteristics

Measuring Value-Added

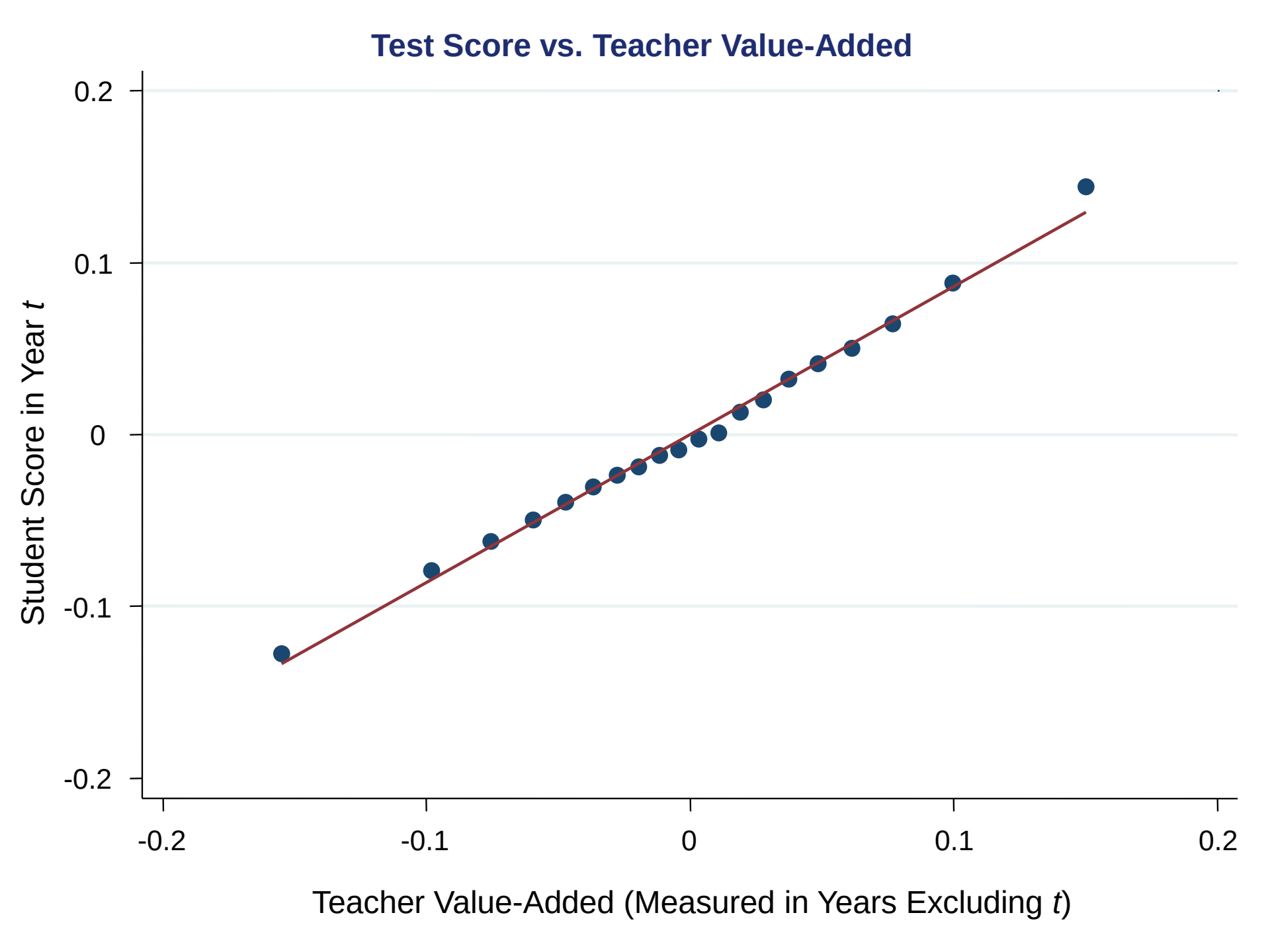
- We measure each teacher's value-added following standard methods used by school districts and researchers [e.g., Kane and Staiger 2008]

- Calculate each teacher's VA in three steps:

  1. Calculate each student's test score *gain* from the year before

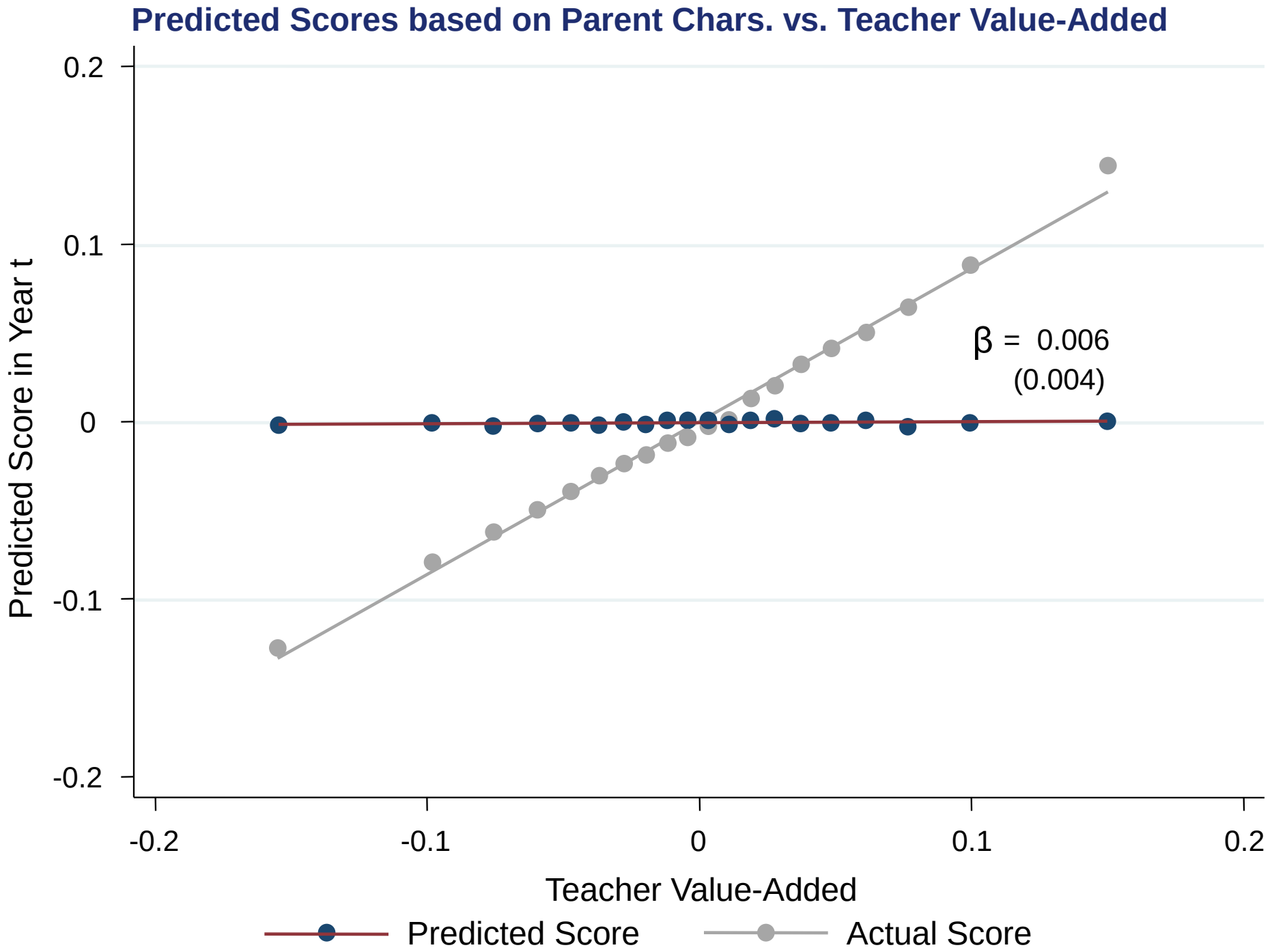
  2. Adjust each student's score gain for differences in characteristics (prior test scores, gender, ESL,...) using a regression model

  3. Compute Teacher VA from average adjusted performance for each student, with a Bayesian shrinkage adjustment for noise


Test Score vs. Teacher Value-Added
0.2
0.1
0
-0.1
-0.2
Student Score in Year t
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added (Measured in Years Excluding t)

Question 1: Are VA Measures Accurate?

- Approach #1: Do higher VA teachers have different types of students?
  - VA models adjusts for some differences using data available in school district records
  - But do students differ based on other characteristics?
  - Test for sorting using data on parent characteristics from tax data
    - Ex: parents' income, marital status, retirement savings, etc.
    - Calculate *predicted* scores based on parent characteristics using OLS regression


Predicted Scores based on Parent Chars. vs. Teacher Value-Added
Predicted Score in Year t
0.2
0.1
0
-0.1
-0.2
β = 0.006
(0.004)
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added
Predicted Score
Actual Score

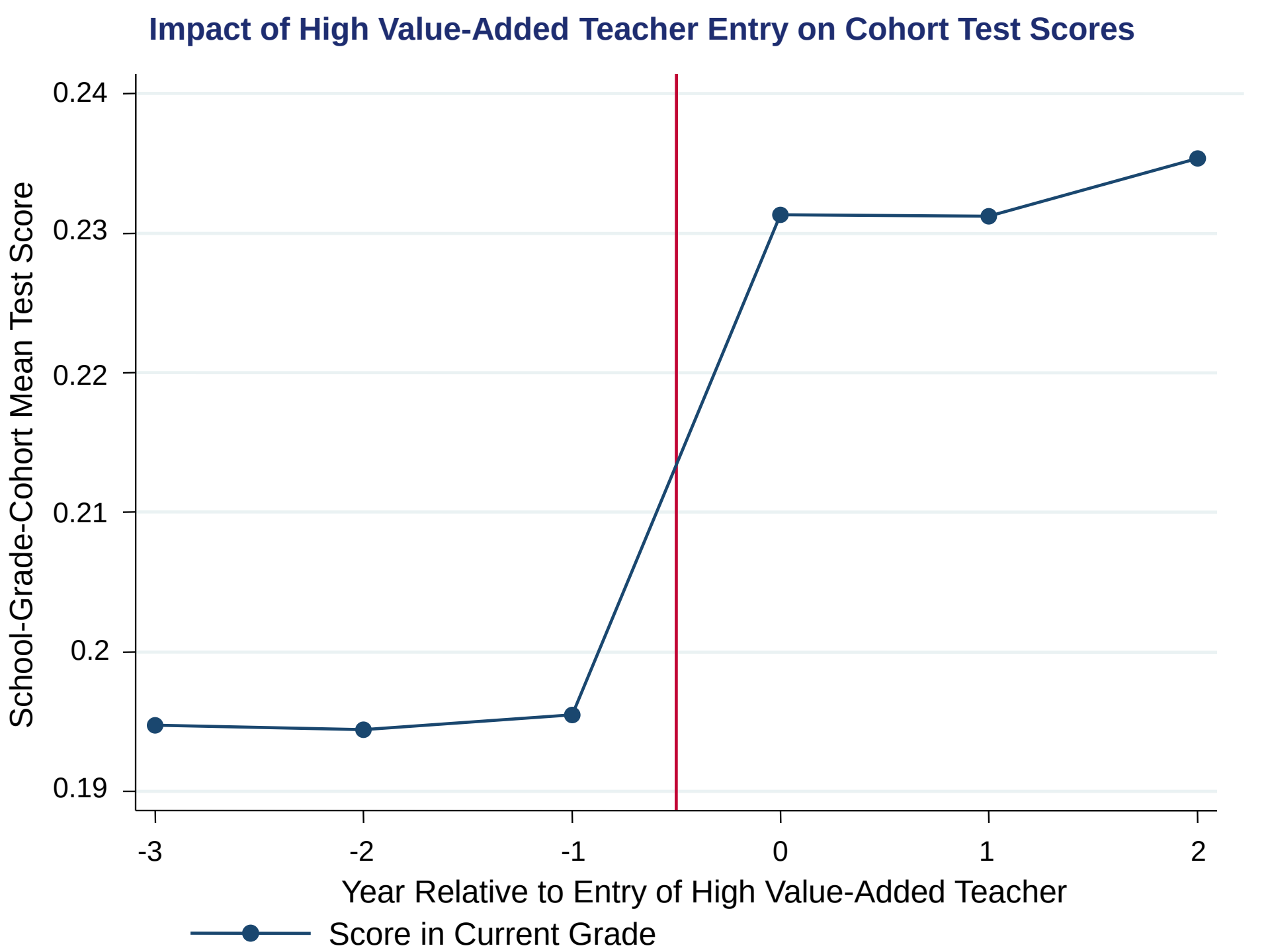
Question 1: Are VA Measures Accurate?

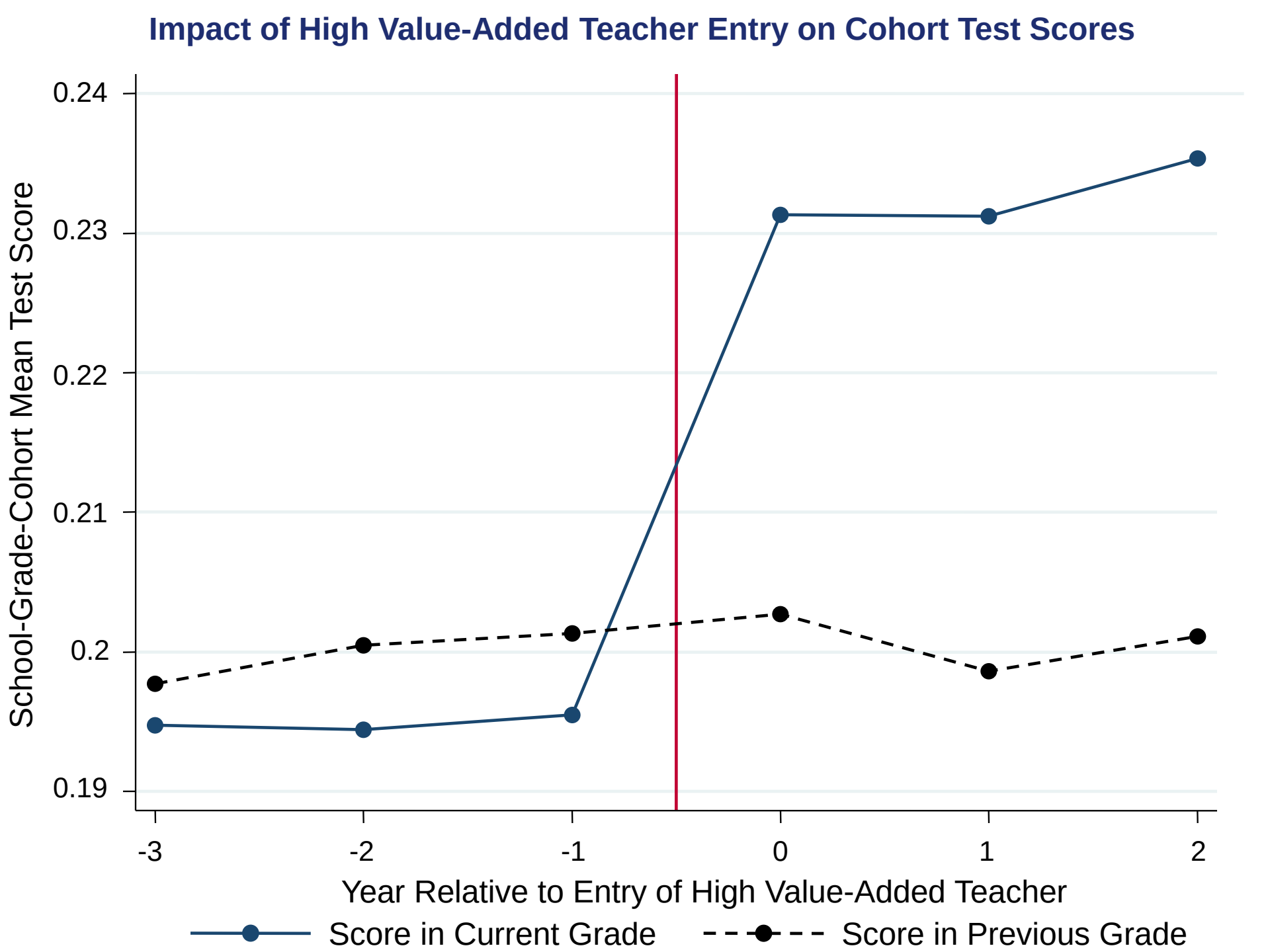
- Approach #2: Quasi-Experimental design based on teacher turnover
  - Ideal experiment: randomly assign students high VA teachers and test if scores go up
  - We use a quasi-experimental approximation to this experiment
    - When high VA teachers arrive at new schools, do scores go up?

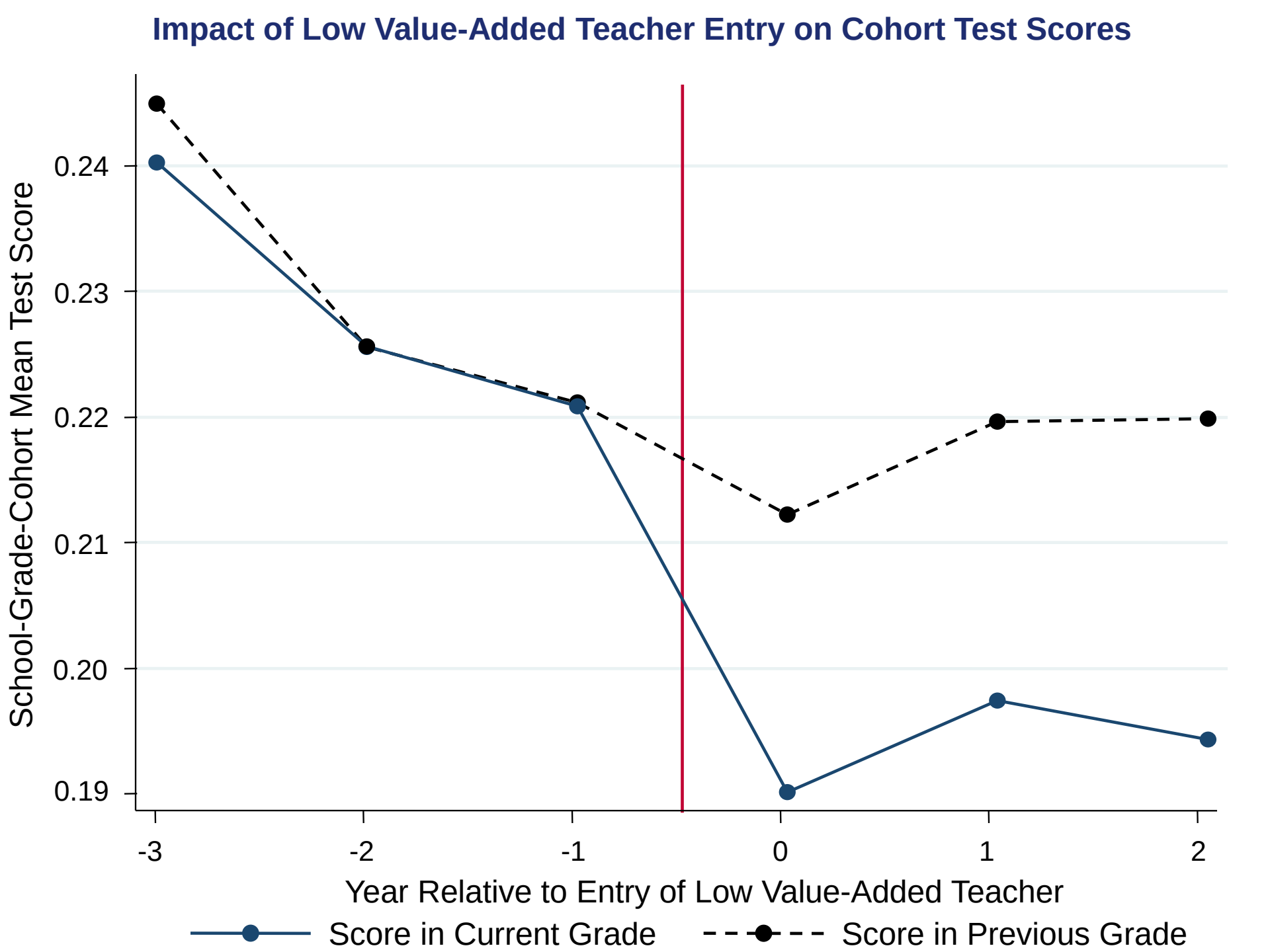
Teacher Switchers in School-Grade-Subject-Year Level Data

School	Grade	Subject	Year	Teachers	Mean Score	Mean Age 28 Earnings
1	5	Math	1992	Smith, Vidoni, …	-.09	$15K
1	5	Math	1993	Smith, Vidoni, …	-.04	$17K
1	5	Math	1994	Smith, Vidoni, …	-.05	$16K
1	5	Math	1995	Ladd, Vidoni, …	0.01	$18K
1	5	Math	1996	Ladd, Vidoni, …	0.04	$17K
1	5	Math	1997	Ladd, Vidoni, …	0.02	$18K

- Smith switches to a different school in 1995; Ladd switches into grade 5


Impact of High Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0.24
0.23
0.22
0.21
0.2
0.19
-3
-2
-1
0
1
2
Year Relative to Entry of High Value-Added Teacher
Score in Current Grade

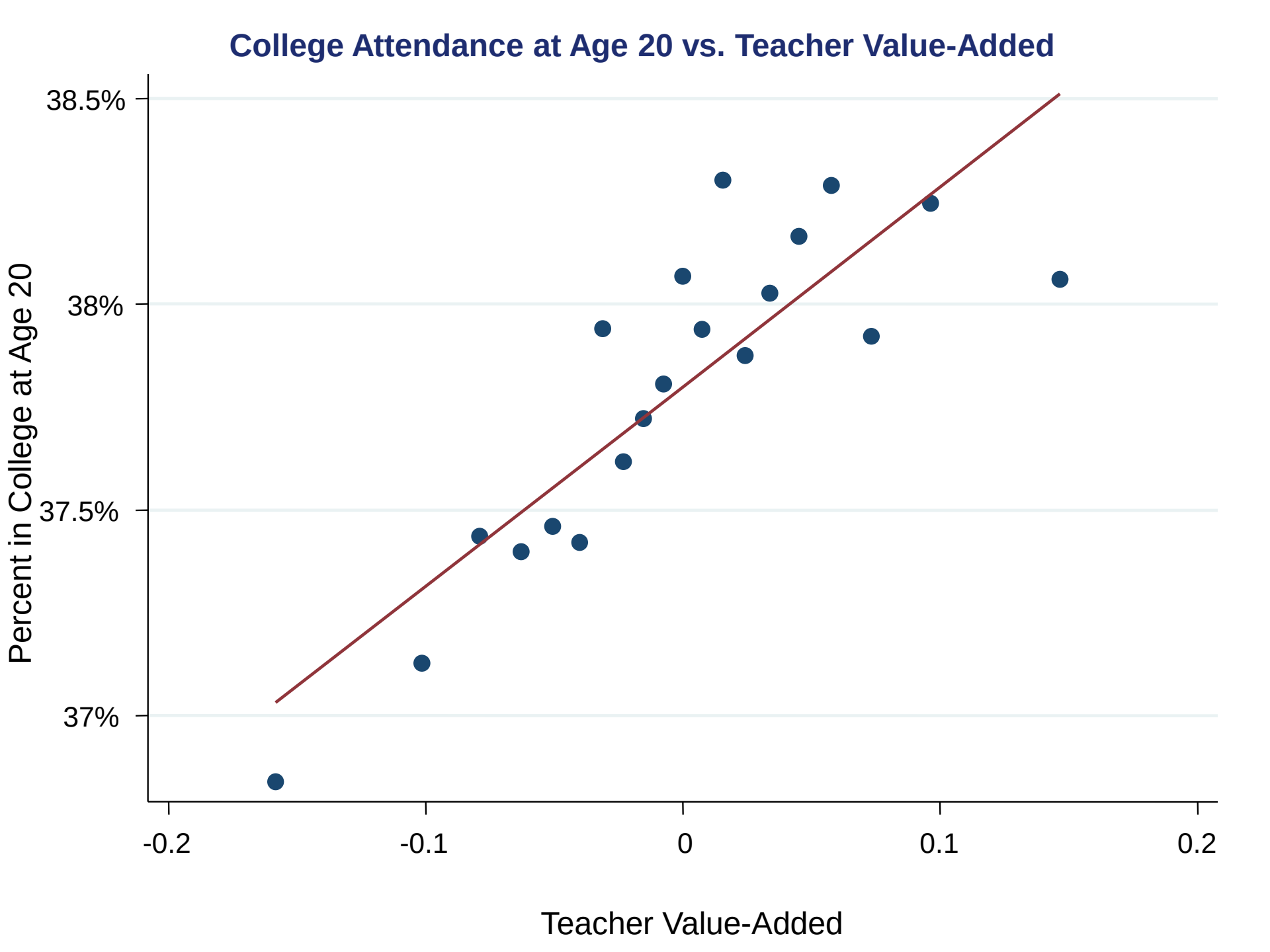

Impact of High Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0.24
0.23
0.22
0.21
0.2
0.19
-3
-2
-1
0
1
2
Year Relative to Entry of High Value-Added Teacher
Score in Current Grade
Score in Previous Grade


Impact of Low Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0.24
0.23
0.22
0.21
0.20
0.19
-3
-2
-1
0
1
2
Year Relative to Entry of Low Value-Added Teacher
Score in Current Grade
Score in Previous Grade

Question 2: Impacts on Outcomes in Adulthood

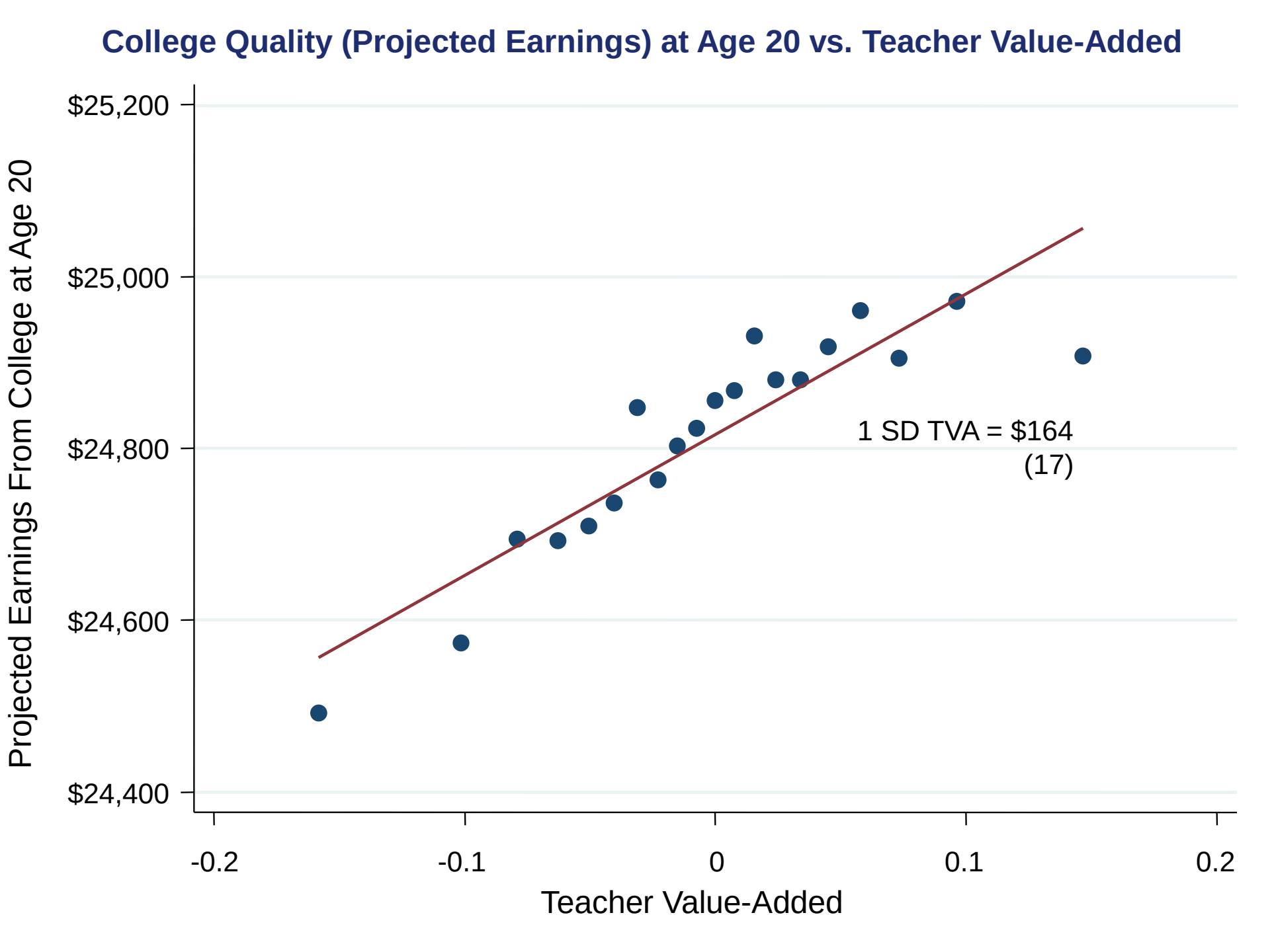
- Now test whether teachers who raise test scores also improve students' long-run outcomes

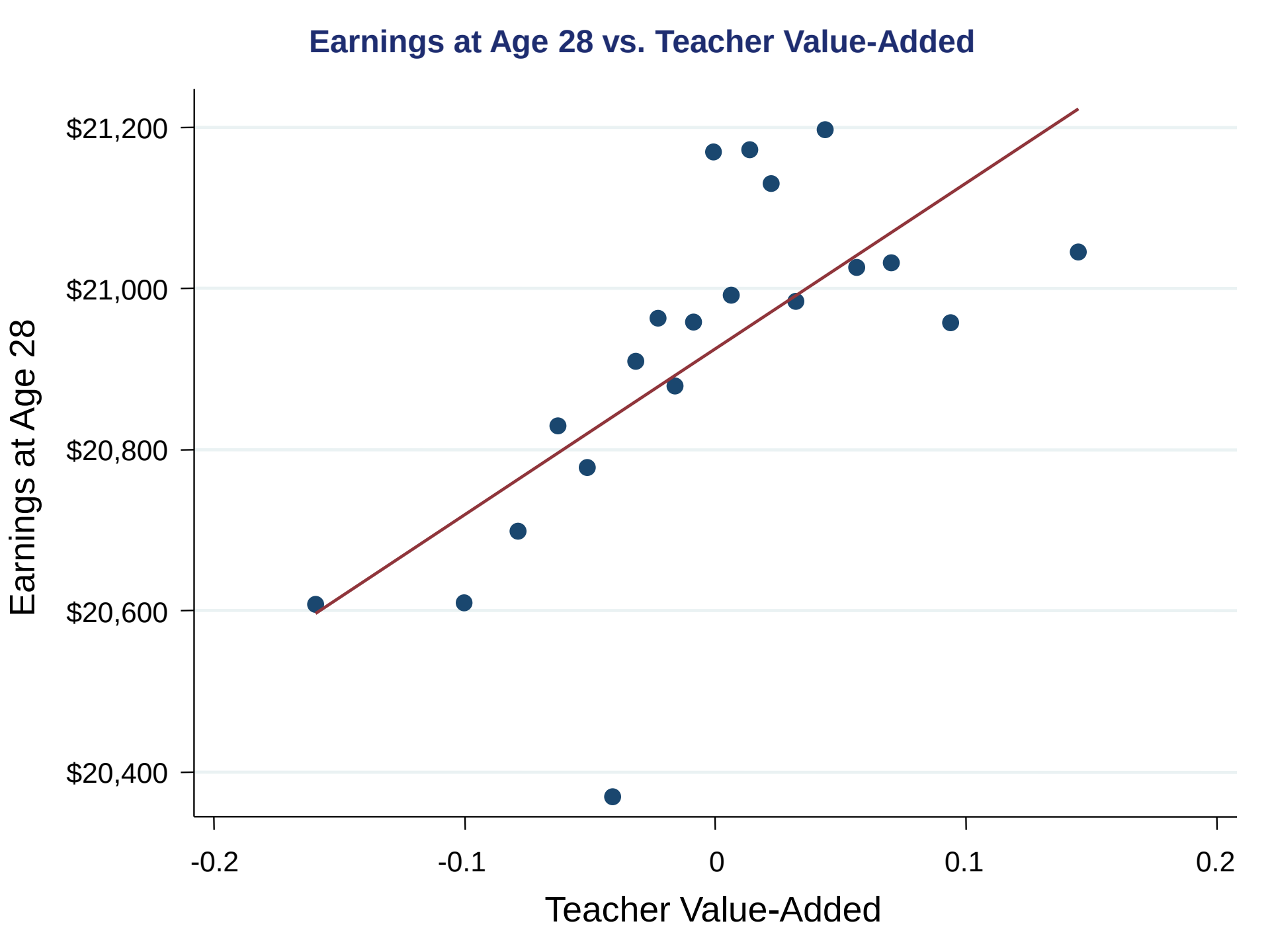
- Interpretation: Impact of having better teacher, as measured by VA, for a **single year** during grades 4-8 on earnings

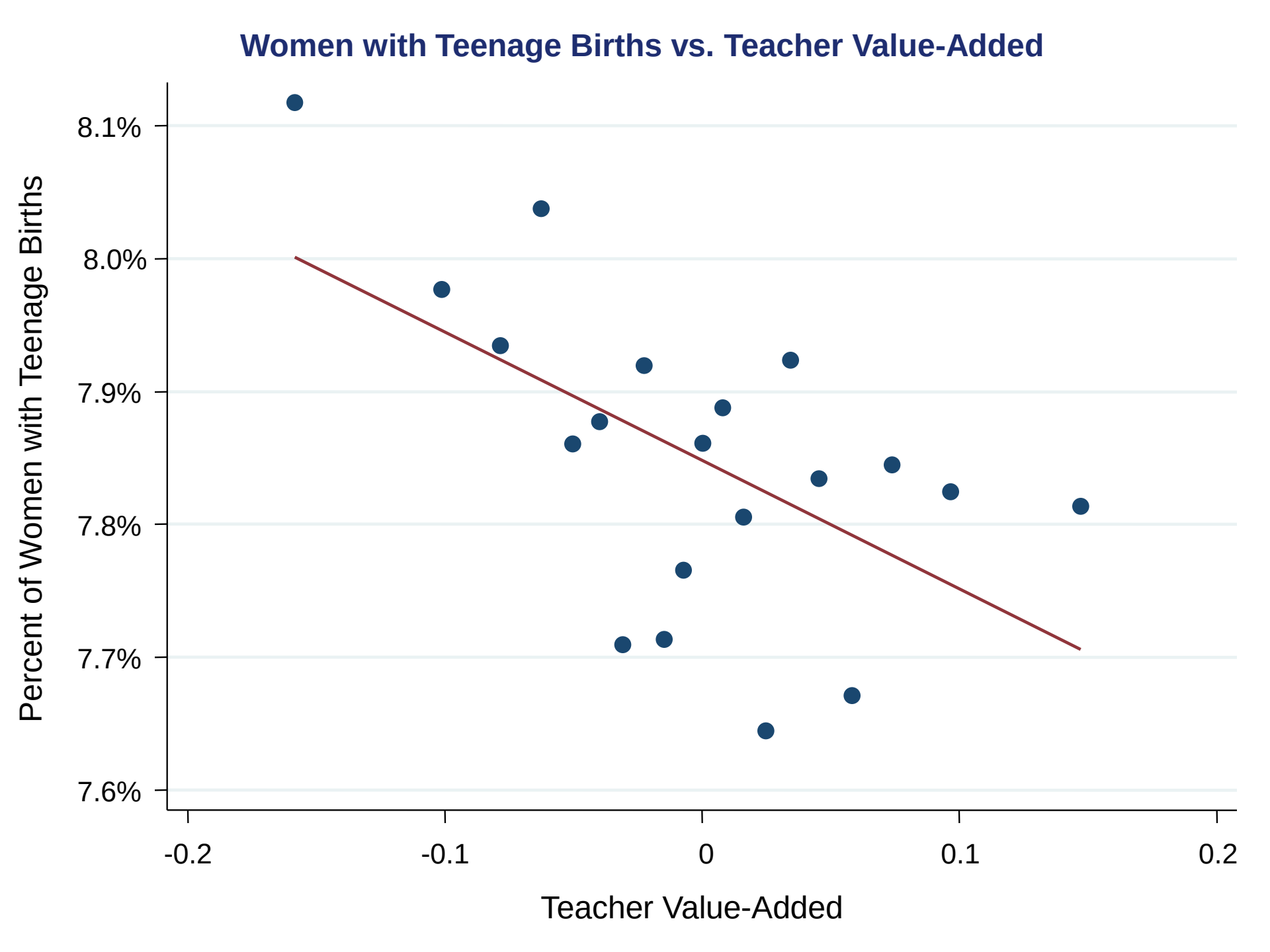

College Attendance at Age 20 vs. Teacher Value-Added
38.5%
38%
37.5%
37%
Percent in College at Age 20
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added

Projected Earnings at Age 30 by College Enrollment at Age 20

US News Ranking	College	Mean Earnings at age 30
1	Harvard	
2	Princeton	
3	Yale	$80,812
4	Cal Tech	
5	MIT	
6	Stanford	
7	U Penn	
8	Columbia	$74,430
9	U Chicago	
10	Duke	
⋮	⋮	
121	Arizona St.	
122	Catholic U	
123	MI Tech	$47,561
124	U Buffalo	
125	U San Fran	
Not in college at age 20		$16,361

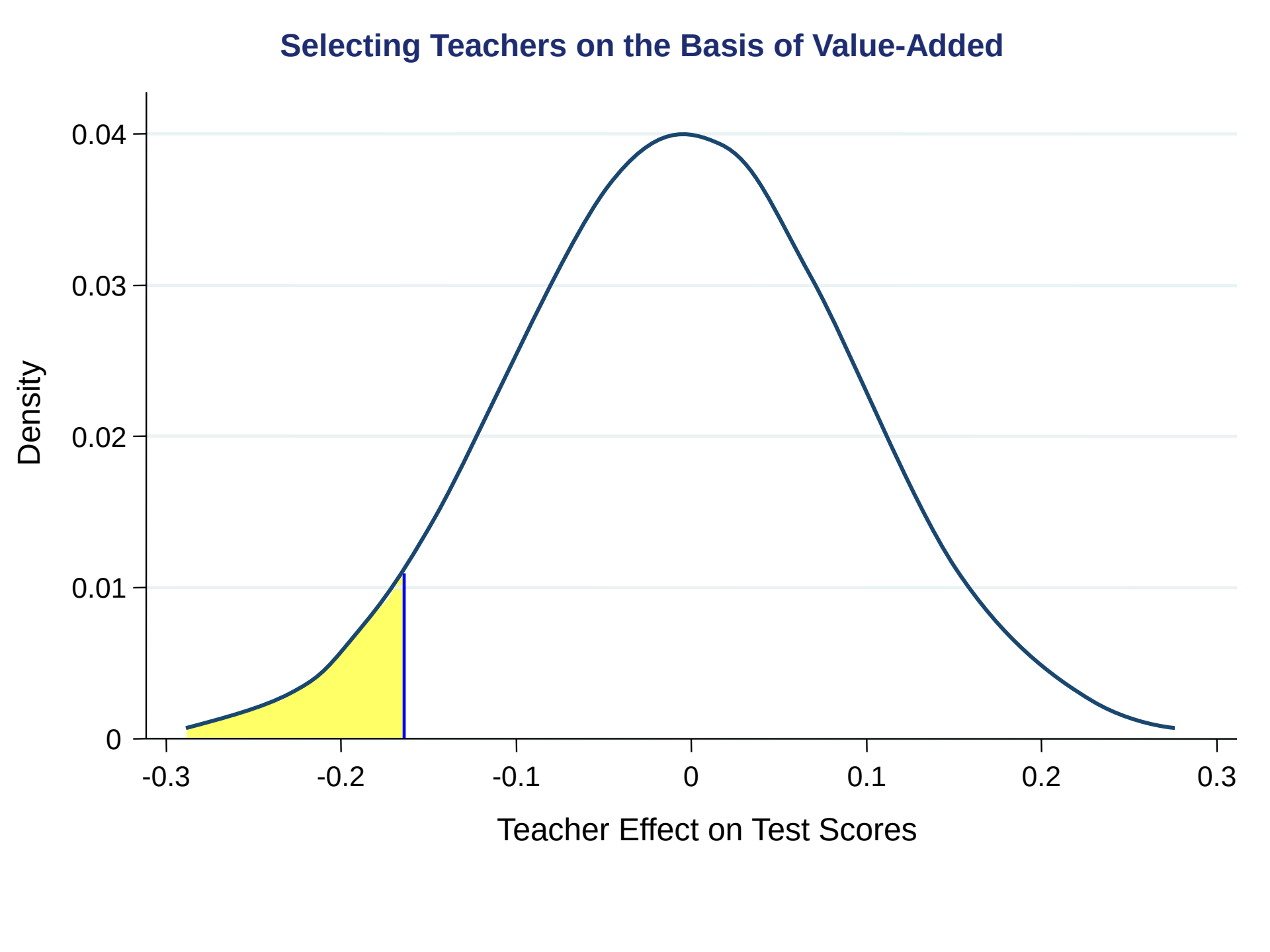

College Quality (Projected Earnings) at Age 20 vs. Teacher Value-Added
$25,200
$25,000
$24,800
$24,600
$24,400
Projected Earnings From College at Age 20
1 SD TVA = $164
(17)
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added


Earnings at Age 28 vs. Teacher Value-Added
$21,200
$21,000
$20,800
$20,600
$20,400
Earnings at Age 28
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added


Women with Teenage Births vs. Teacher Value-Added
Percent of Women with Teenage Births
8.1%
8.0%
7.9%
7.8%
7.7%
7.6%
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added

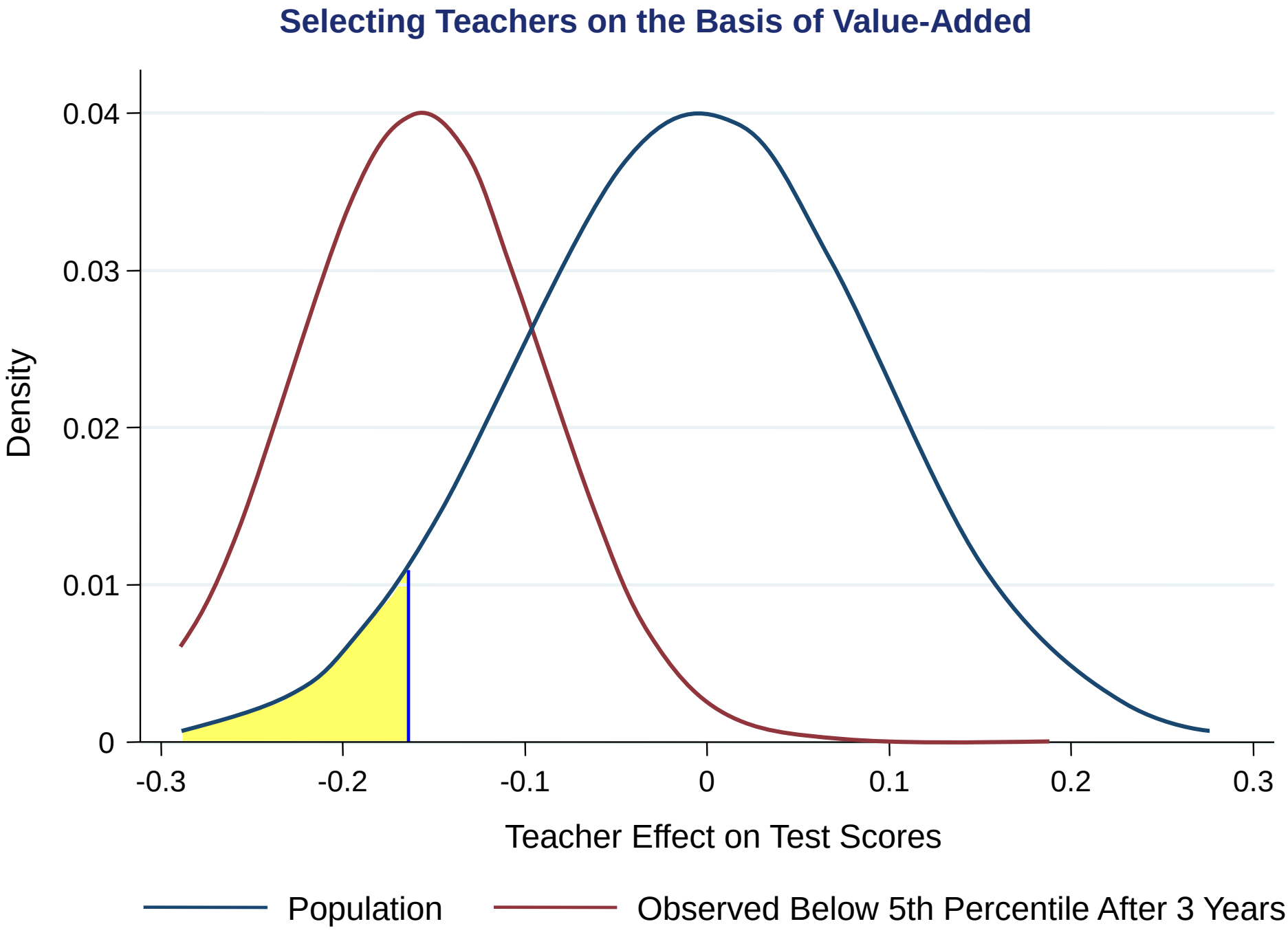
Question 3: Stability of VA and Policy Relevance

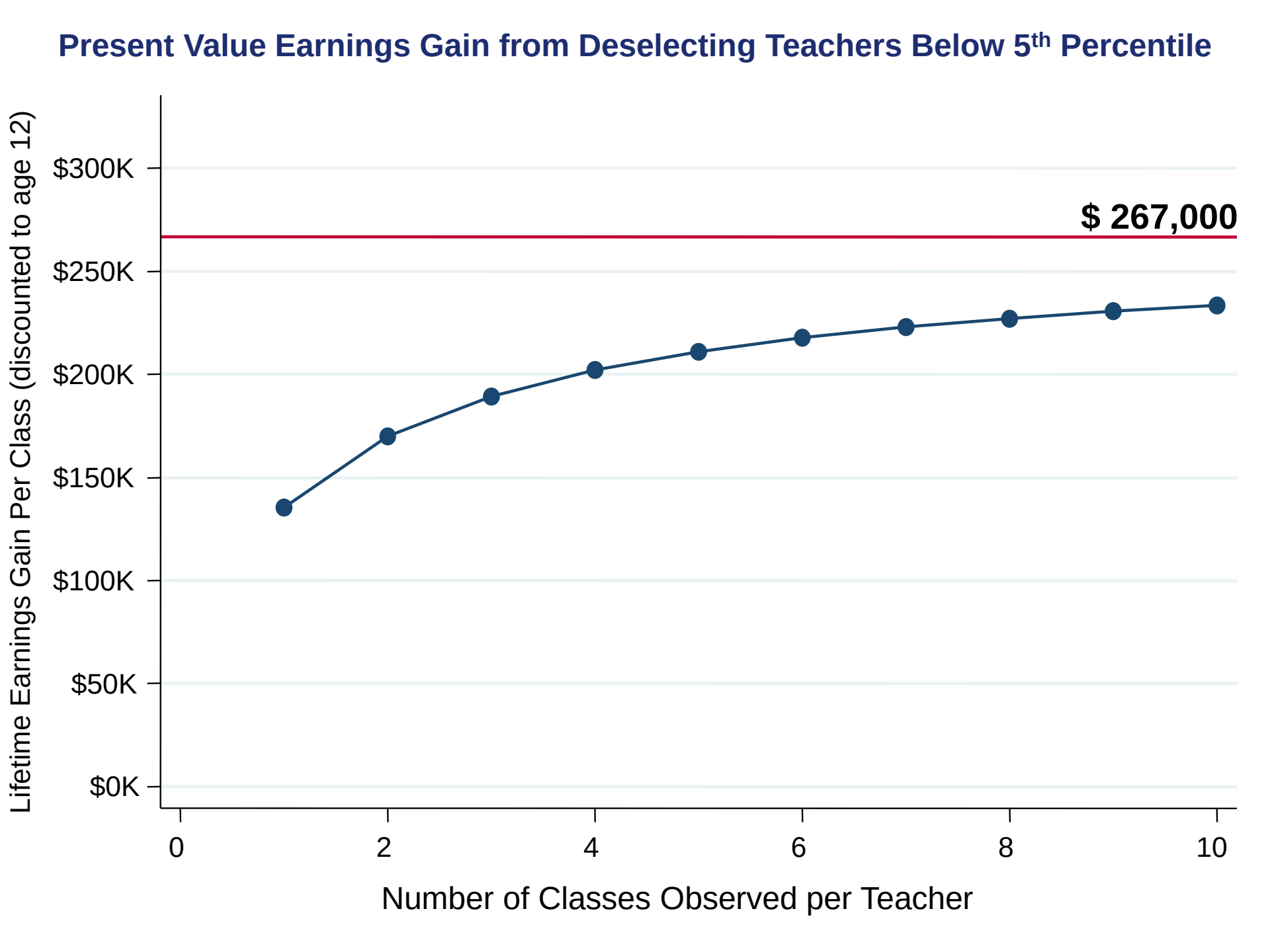
- Any evaluation of teachers based on VA must rely on only a few years of classroom data
  - This generates noise in VA estimates, potentially reducing its utility for performance evaluation
- Evaluating magnitude of noise requires a policy-relevant metric
  - Frequently-cited correlation coefficients and measures of stability across years not directly informative
- As an illustration, we analyze impacts of selecting teachers based on their VA [Hanushek 2009, Rothstein 2012]


Selecting Teachers on the Basis of Value-Added
0.04
0.03
0.02
0.01
0
Density
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Effect on Test Scores

The Value of Good Teachers

- On average, replacing a bottom 5% teacher with an average teacher for one year raises a child's cumulative lifetime income by $50,000
    - For a class of average size (28 students), cumulative lifetime income gains from a high VA teacher surpass $1.4 million
- Equivalent to $267,000 in present value at age 12, discounting future earnings gains at a 5% interest rate
- Note that selection is not the only policy tool to achieve these gains: raising a given teacher's VA would in principle have the same impact


Selecting Teachers on the Basis of Value-Added
0.04
0.03
0.02
0.01
0
Density
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Effect on Test Scores
Population
Observed Below 5th Percentile After 3 Years

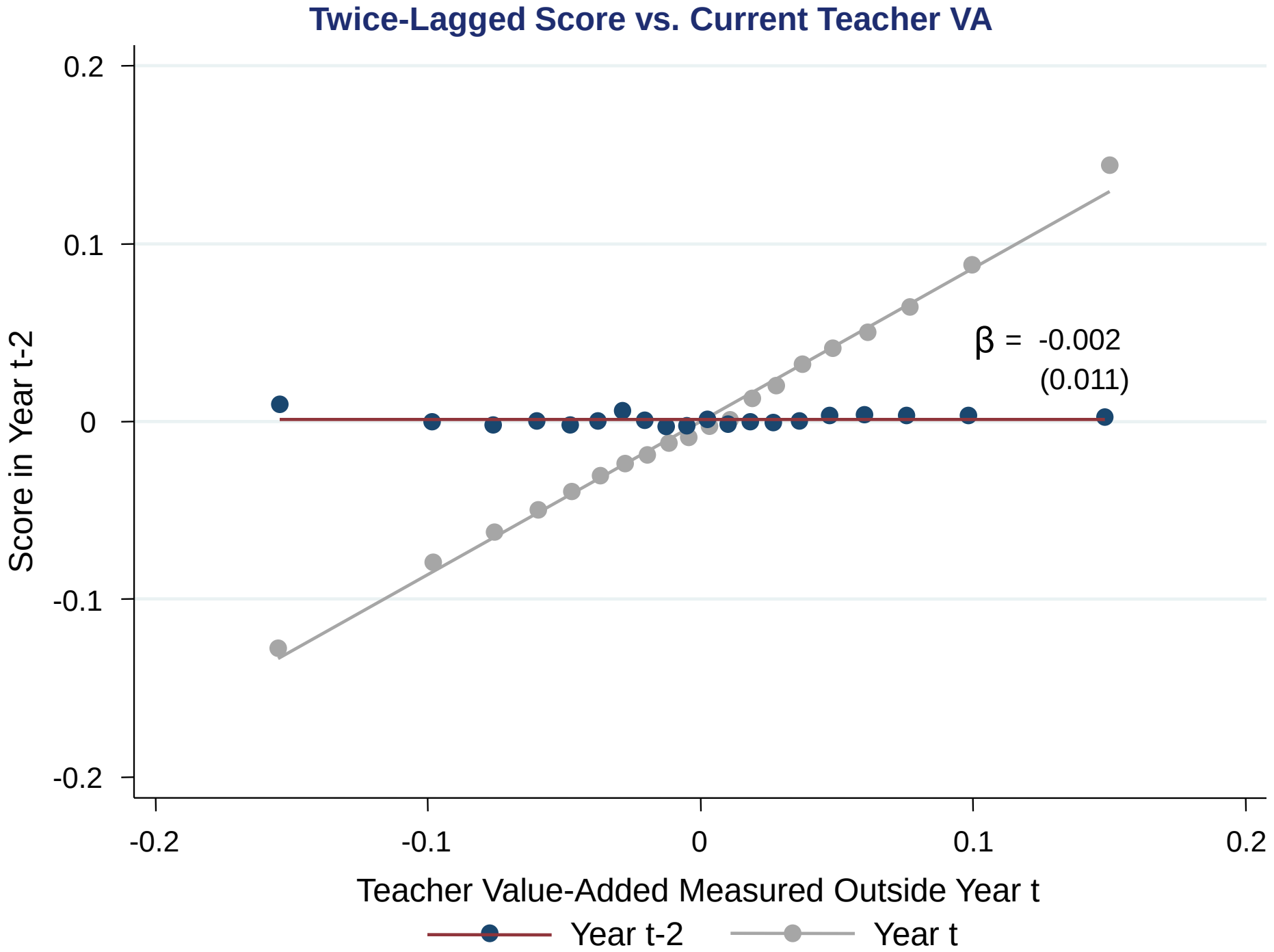

Present Value Earnings Gain from Deselecting Teachers Below 5th Percentile
Lifetime Earnings Gain Per Class (discounted to age 12)
$300K
$250K
$200K
$150K
$100K
$50K
$0K
$ 267,000
0
2
4
6
8
10
Number of Classes Observed per Teacher

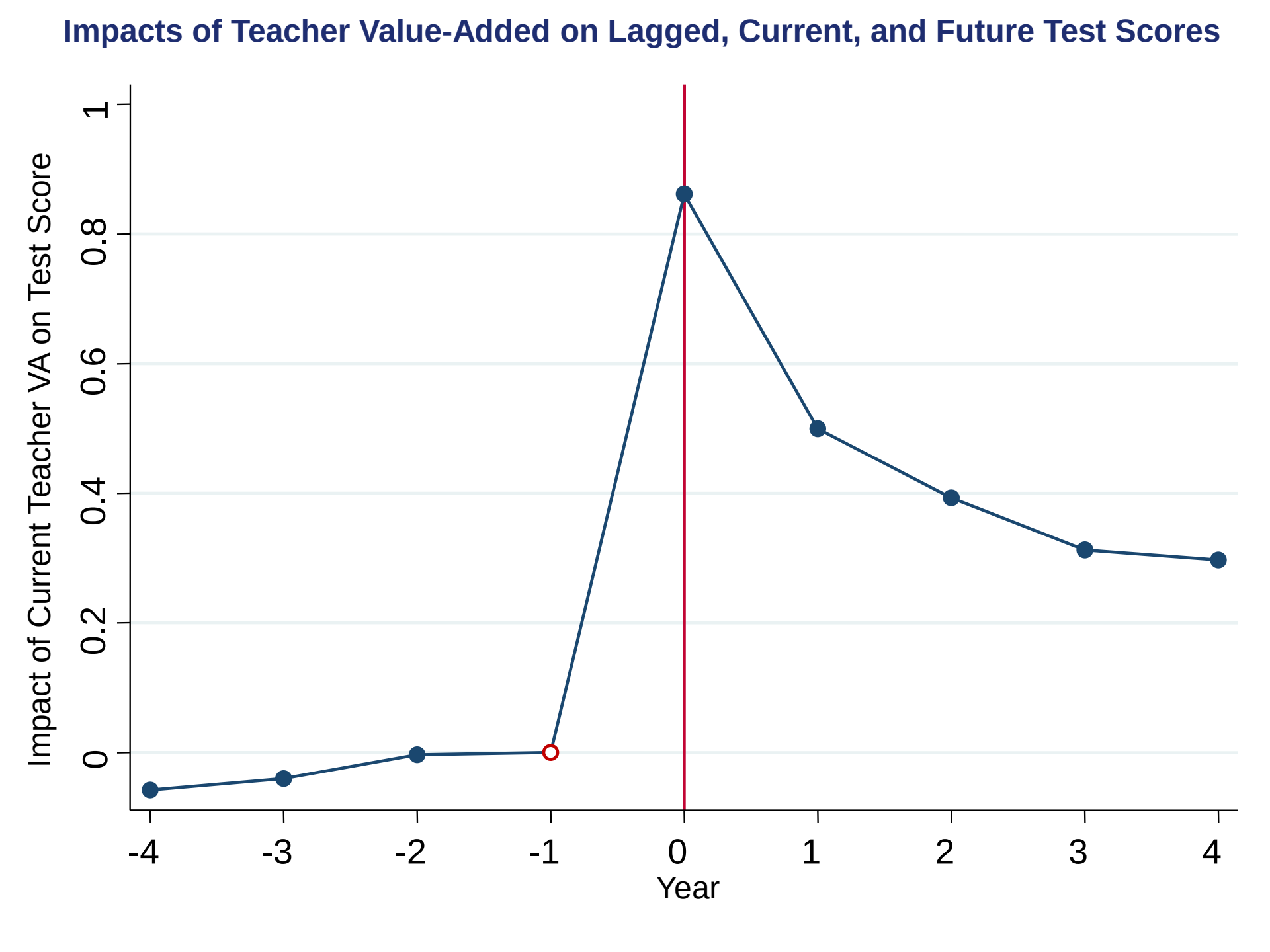
Using Value-Added for Teacher Evaluation

- Test scores can provide *one* useful input into teacher evaluation
- But further work is needed to assess VA as a policy tool
- Most important concern: behavioral responses to high-stakes testing [Barlevy and Neal 2011]
  - Using VA measures in high-stakes evaluation could induce negative behavioral responses such as teaching to the test or cheating
  - Can only address this issue empirically by studying districts where VA is starting to be used (e.g., Washington DC)

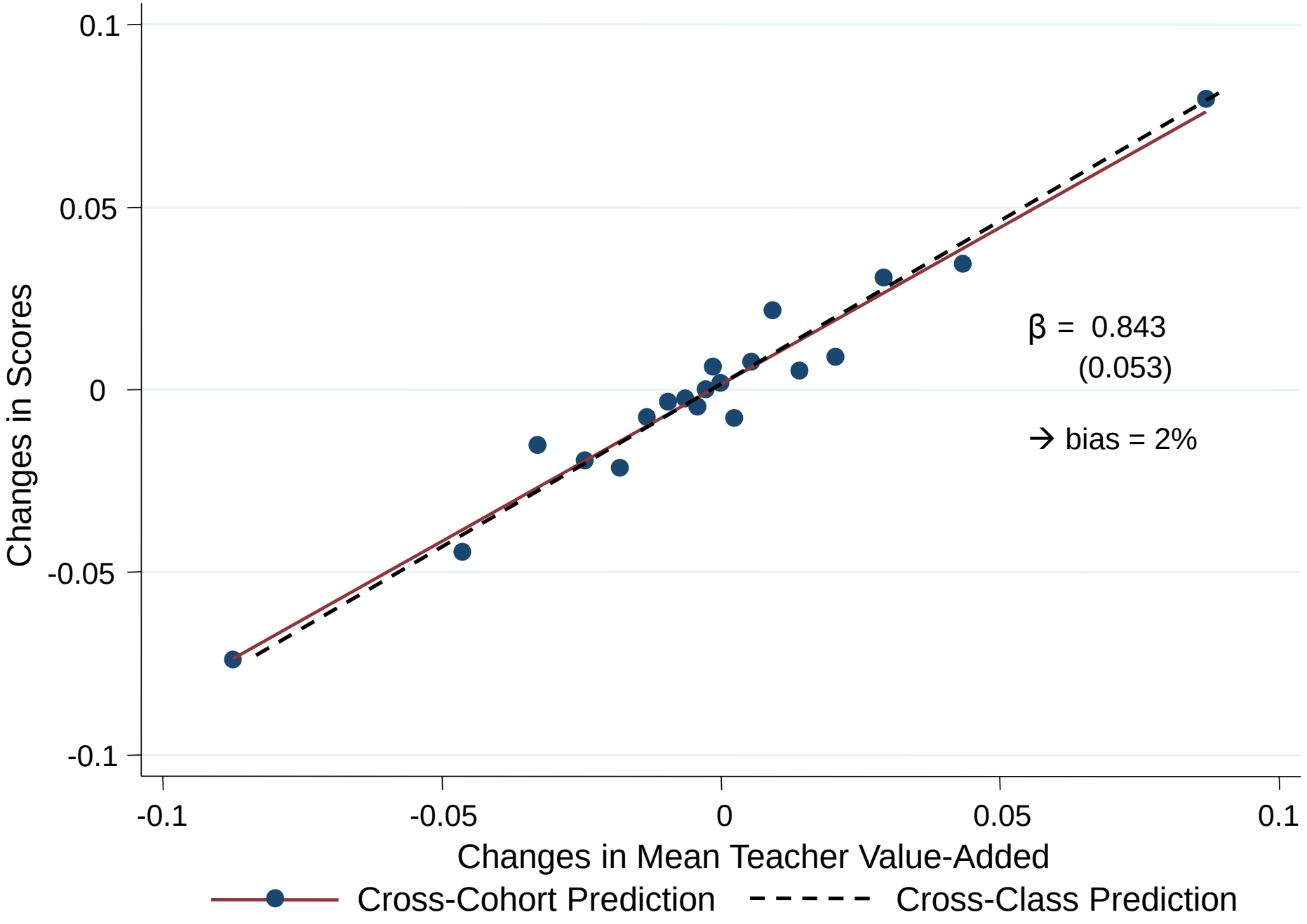
→ Main lesson of present study: large potential returns from improving teacher quality, whether using VA or other tools

Slides to Answer Questions


Twice-Lagged Score vs. Current Teacher VA
0.2
0.1
0
-0.1
-0.2
Score in Year t-2
β = -0.002
(0.011)
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added Measured Outside Year t
Year t-2
Year t

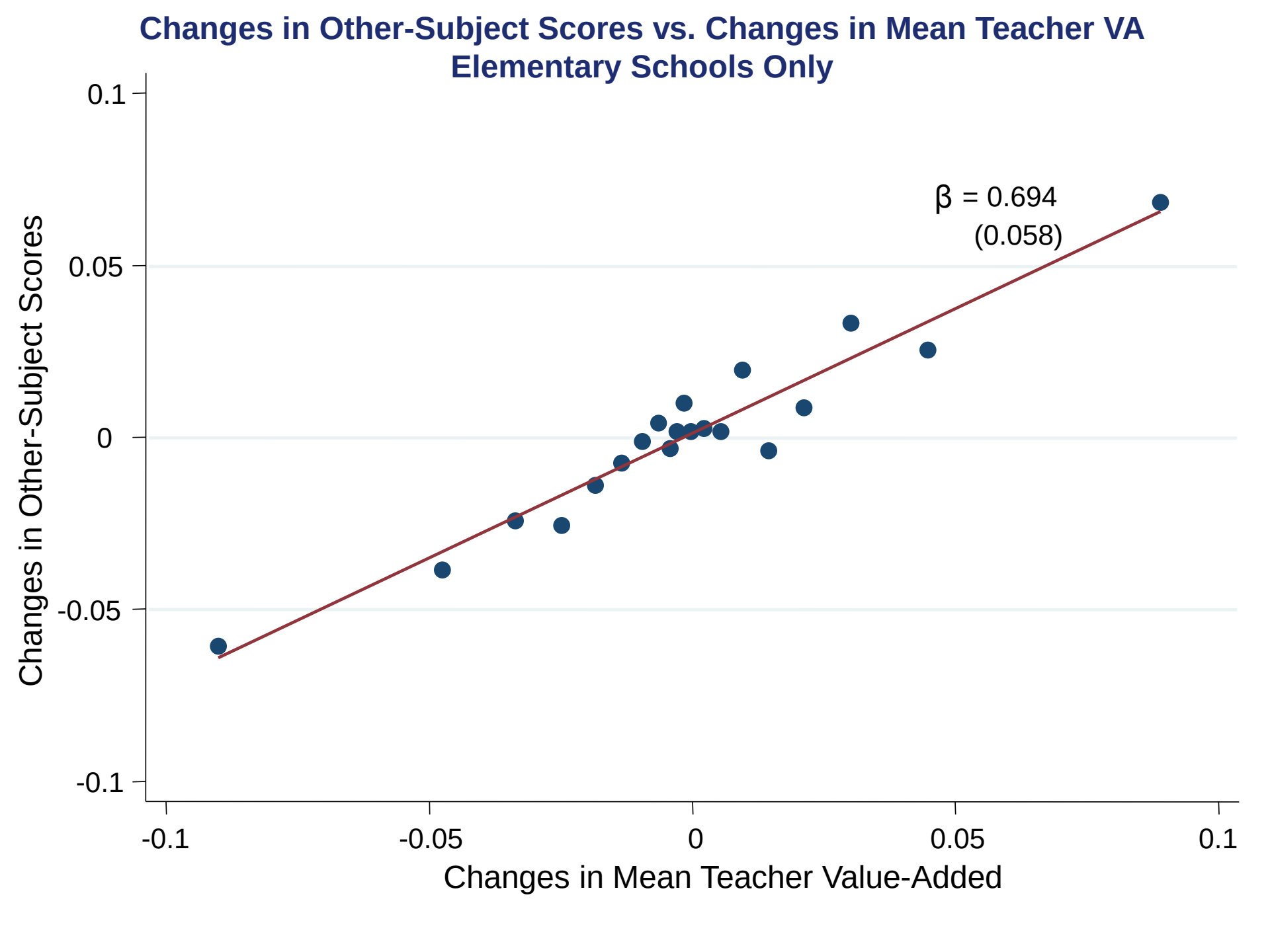

Impacts of Teacher Value-Added on Lagged, Current, and Future Test Scores
Impact of Current Teacher VA on Test Score
0
0.2
0.4
0.6
0.8
1
-4
-3
-2
-1
0
1
2
3
4
Year

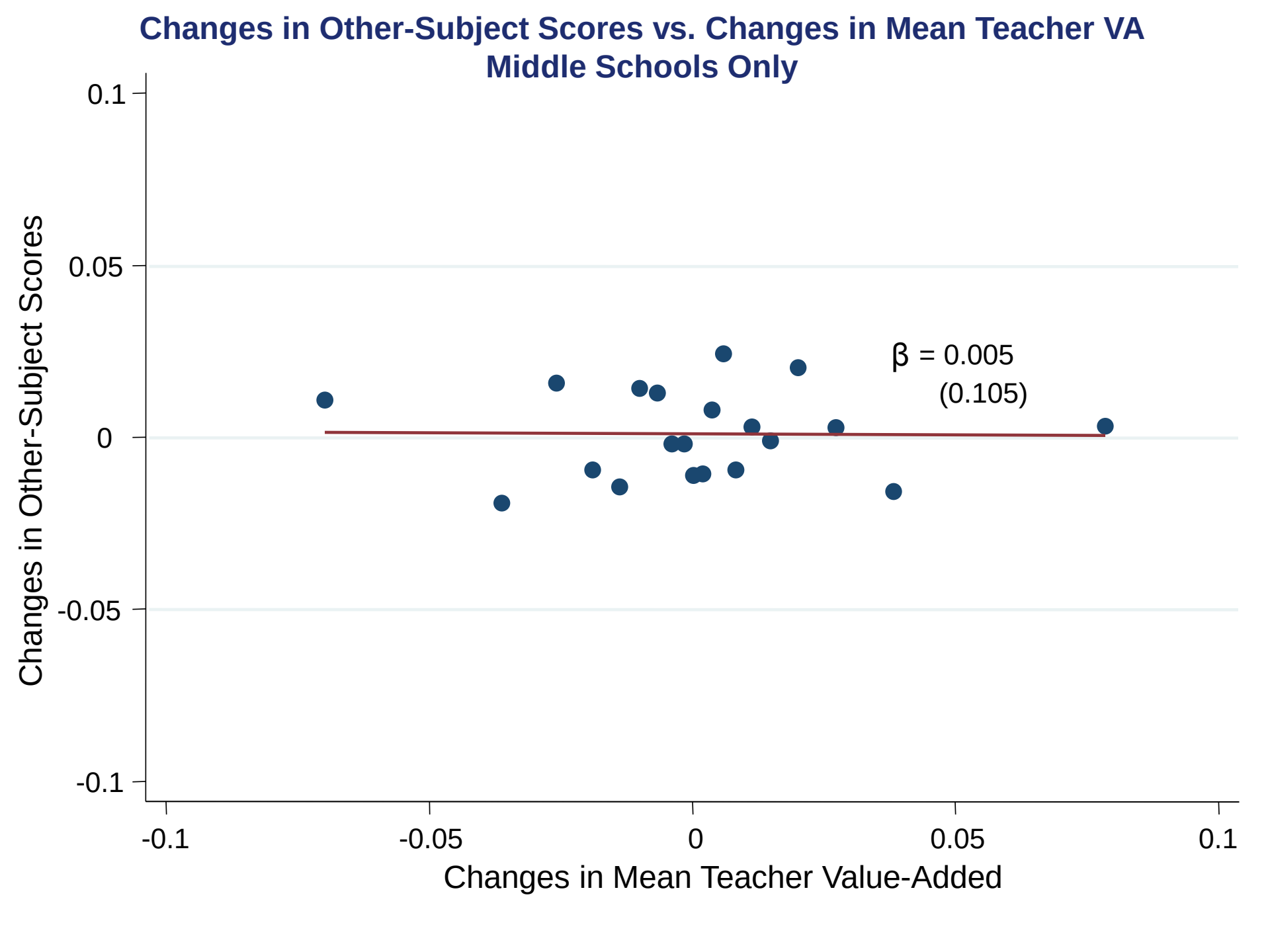
Teacher Switchers Design: Changes in Scores vs. Changes in Mean Teacher VA

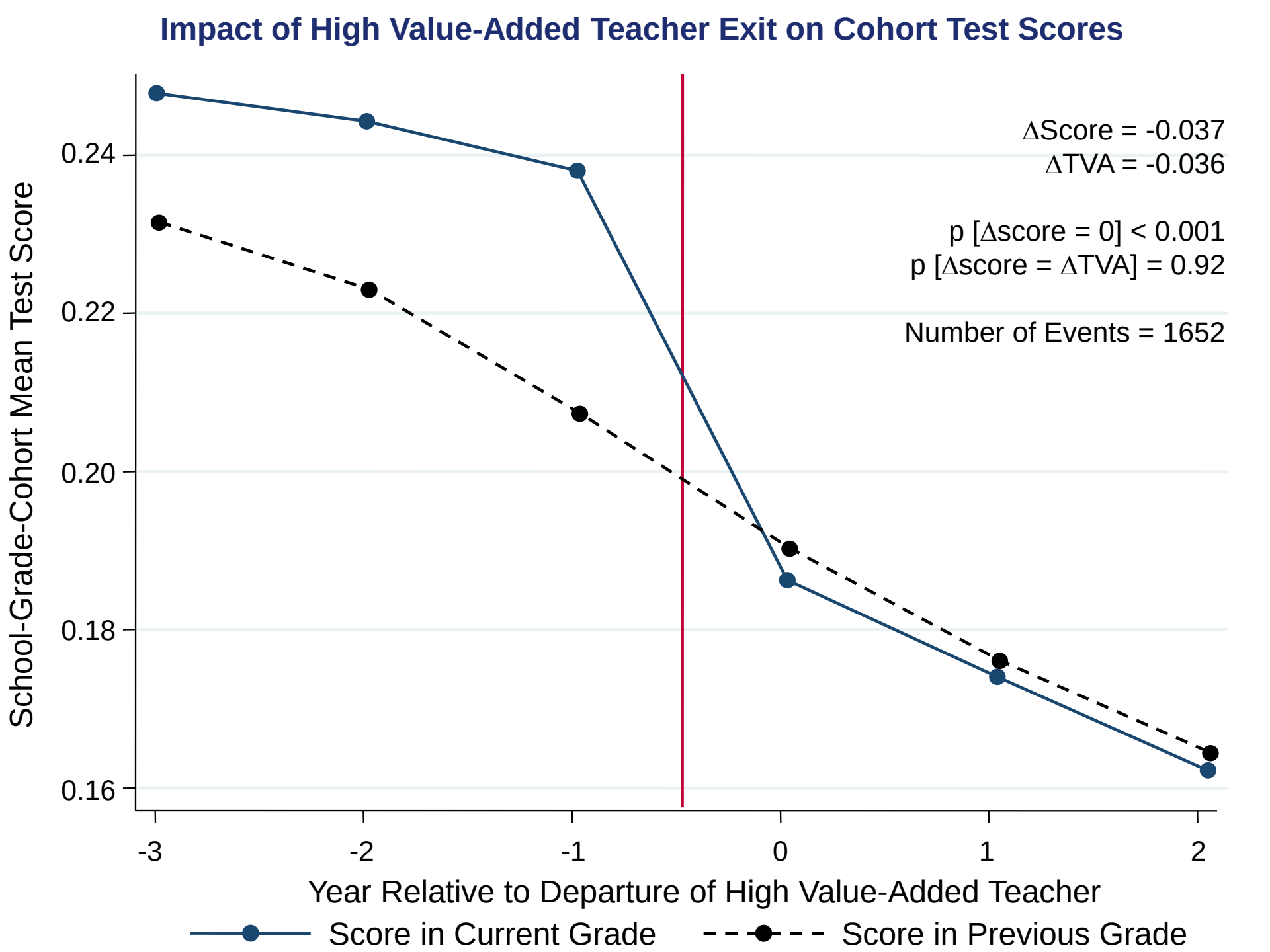


Sensitivity of Teacher Value-Added Measures to Controls

	(1) baseline	(2) add parent chars.	(3) add t-2 scores	(4) t-1 scores only	(5) no controls	(6) Quasi-Experimental Estimate of Bias
Baseline	1.000					3.1% (7.6)
add Parent	0.999	1.000				2.6% (7.6)
add t-2 Scores	0.975	0.974	1.000			1.7% (7.4)
t-1 Scores only	0.945	0.943	0.921	1.000		**14.3%** (6.9)
No Controls	0.296	0.292	0.279	0.323	1.000	87.8% (1.4)


Changes in Other-Subject Scores vs. Changes in Mean Teacher VA
Elementary Schools Only
β = 0.694
(0.058)
0.1
0.05
0
-0.05
-0.1
Changes in Other-Subject Scores
-0.1
-0.05
0
0.05
0.1
Changes in Mean Teacher Value-Added


Changes in Other-Subject Scores vs. Changes in Mean Teacher VA
Middle Schools Only
β = 0.005
(0.105)
0.1
0.05
0
-0.05
-0.1
Changes in Other-Subject Scores
-0.1
-0.05
0
0.05
0.1
Changes in Mean Teacher Value-Added


Impact of High Value-Added Teacher Exit on Cohort Test Scores
ΔScore = -0.037
ΔTVA = -0.036
p [Δscore = 0] < 0.001
p [Δscore = ΔTVA] = 0.92
Number of Events = 1652
School-Grade-Cohort Mean Test Score
0.24
0.22
0.20
0.18
0.16
-3
-2
-1
0
1
2
Year Relative to Departure of High Value-Added Teacher
Score in Current Grade
Score in Previous Grade