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

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

March 2014

Introduction: Teacher Value-Added

- How can we measure and improve the quality of teaching in elementary schools?

- One approach: "value-added" (VA) measures [Hanushek 1971, Murnane 1975, Rockoff 2004, Rivkin et al. 2005, Aaronson et al. 2007]
  - Rate teachers based on their students' test score gains

- School districts have started to use VA to evaluate teachers, leading to considerable debate
  - Ex: Washington D.C. lays off teachers and offers bonuses using a metric that puts 50% weight on VA measures
  - Lawsuit in LA based on VA measures

Debate About Teacher Value-Added

- Debate stems primarily from two intellectual issues:

  1. Disagreement about whether VA measures are biased
     [Kane and Staiger 2008, Rothstein 2010]
     - Do differences in test-score gains across teachers capture causal impacts of teachers or are they driven by student sorting?
     - If VA estimates are biased, they will incorrectly reward or penalize teachers for the mix of students they get

  2. Lack of evidence on teachers' long-term impacts
     - Do teachers who raise test scores improve students' long-term outcomes or are they simply better at teaching to the test?

Objectives of This Project

- This study answers these two questions by tracking one million children from childhood to early adulthood
  - Develop new quasi-experimental tests for bias in VA estimates
  - Test whether children who get high VA teachers have better outcomes in adulthood

- Results also shed light on broader issues in the economics of education
  - What are the long-run returns to investments in better teaching?
  - Are impacts on scores a good proxy for long-term impacts of educational interventions?

Outline

1. Data
2. Construction of Value-Added Estimates with Drift
3. Evaluating Bias in Value-Added Estimates
4. Long-Term Impacts
5. Policy Implications

Dataset 1: School District Data

- Teacher and class assignments from 1991-2009 for 2.5 million children
- Test scores from 1989-2009
  - Scaled scores standardized by grade and subject (math/reading)
  - 18 million test scores, grades 3-8
- Exclude students in special ed. schools and classrooms (6% of obs.)

Dataset 2: United States Tax Data

- Selected data from U.S. federal income tax returns from 1996-2010
  - Includes non-filers via information forms (e.g. W-2's)
- Student outcomes: earnings, college, teenage birth, neighborhood quality
- Parent characteristics: household income, 401k savings, home ownership, marital status, age at child birth
  - Omitted variables from standard VA models
- Approximately 90% of student records matched to tax data
  - Data were analyzed as part of a broader project on tax policy
  - Research based purely on statistics aggregating over thousands of individuals, not on individual data

Data Structure

Student	Subject	Year	Grade	Class	Teacher	Test Score	Age 28 Earnings
			⋮				
Raj	Math	1992	4	1	Samuelson	0.5	$22K
Raj	English	1992	4	1	Samuelson	1.3	$22K
Raj	Math	1993	5	2	Solow	0.9	$22K
Raj	English	1993	5	2	Solow	0.1	$22K
Raj	Math	1994	6	3	Arrow	1.5	$22K
Raj	English	1994	6	4	Stigler	0.5	$22K
			⋮				

- One observation per student-subject-year

Summary Statistics

Variable	Mean (1)	S.D. (2)
Student Data:		
Class size (not student-weighted)	28.2	5.8
Test score (SD)	0.12	0.91
Female	50.4%	
Age (years)	11.7	1.6
Free lunch eligible (1999-2009)	77.1%	
Minority (Black or Hispanic)	72.1%	
English language learner	4.9%	
Special education	3.1%	
Repeating grade	2.7%	
Number of subject-school years per student	6.25	3.18
Student match rate to adult outcomes	89.2%	
Student match rate to parent chars.	94.8%	

Summary Statistics

Variable	Mean (1)	S.D. (2)
Adult Outcomes:		
Annual wage earnings at age 20	5,670	7,773
Annual wage earnings at age 25	17,194	19,889
Annual wage earnings at age 28	20,885	24,297
In college at age 20	35.6%	
In college at age 25	16.5%	
College Quality at age 20	26,408	13,461
Contribute to a 401(k) at age 25	19.1%	
ZIP code % college graduates at age 25	13.7%	
Had a child while a teenager (for women)	14.3%	
Parent Characteristics:		
Household income (child age 19-21)	40,808	34,515
Ever owned a house (child age 19-21)	34.8%	
Contributed to a 401k (child age 19-21)	31.3%	
Ever married (child age 19-21)	42.2%	
Age at child birth	28.3	7.8
Predicted Score	0.17	0.26

Constructing Value-Added Estimates

- Simplest case: teachers teach one class per year with N students
- All teachers have test score data available for t previous years
- Objective: predict test scores for students taught by teacher j in year $t+1$ using test score data from previous t years

Constructing Value-Added Estimates

- Three steps to estimate VA in year $t+1$

  1. Form residual test scores, controlling for observables

     - Regress test scores A_{is} on observable student characteristics X_{is}, including prior test scores $A_{i,s-1}$ using <u>within-teacher</u> variation

  2. Regress mean class-level test score residuals in year t on class-level test score residuals in years 0 to $t-1$

  3. Use estimated coefficients $\psi_1, \dots, \psi_t$ to predict VA in year $t+1$ based on mean test score residuals in years 1 to t for each teacher j

- Paper generalizes this approach to allow for variation in numbers of students and classes across teachers

Constructing Value-Added Estimates

- Practical complications: number of students varies across classes, number of years varies across teachers, multiple classes per year in middle school
- Generalize regression approach by estimating an autocorrelation vector and assume stationarity of teacher VA process
- Then form a prediction for VA in each teacher-year using data from all *other* years using autocorrelation vector
- STATA ado file to implement this procedure on the web

Constructing Value-Added: Special Cases

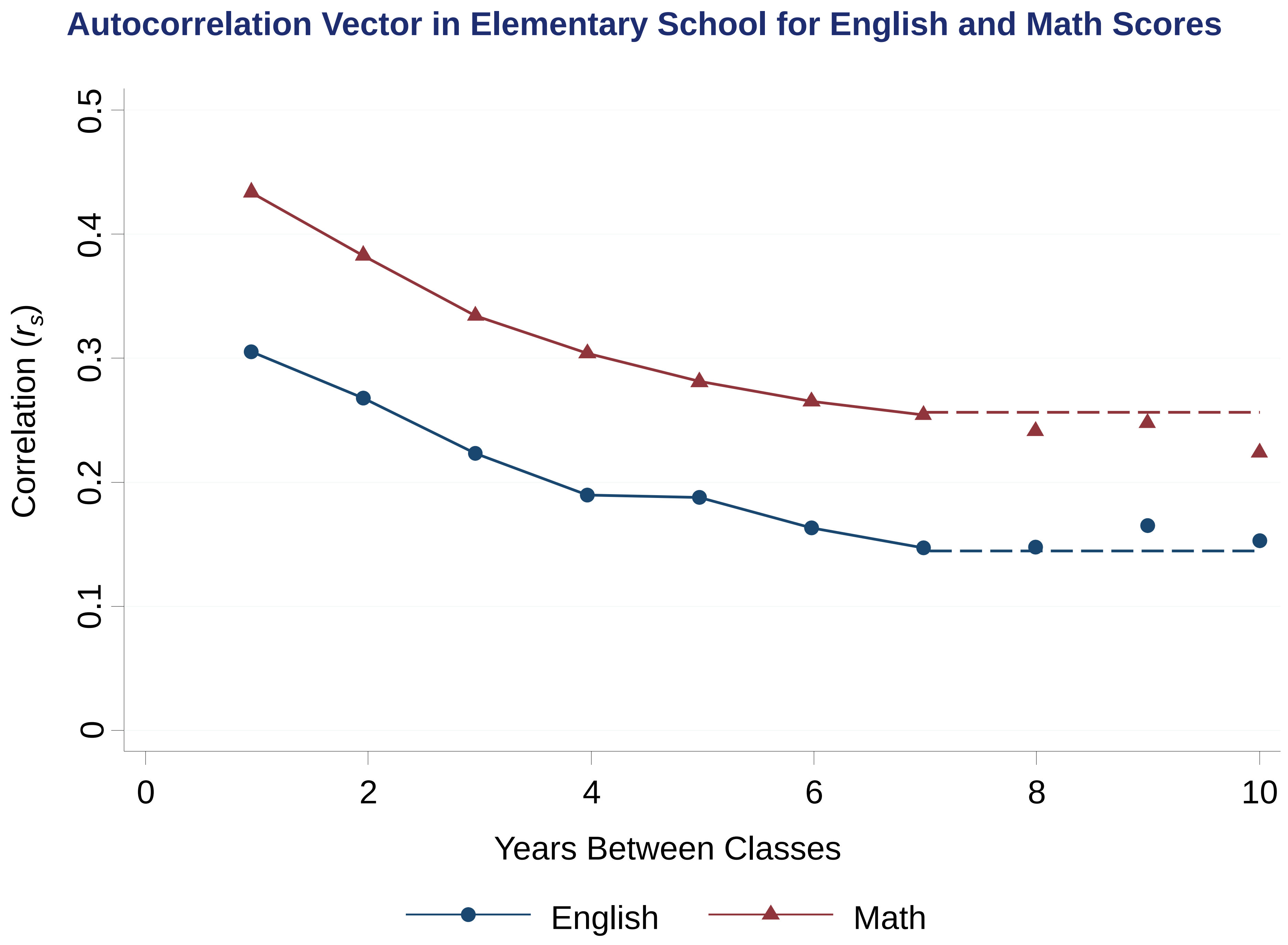
- Two special cases:

1. Forecast VA in year *t* using data from only year *t-s*:

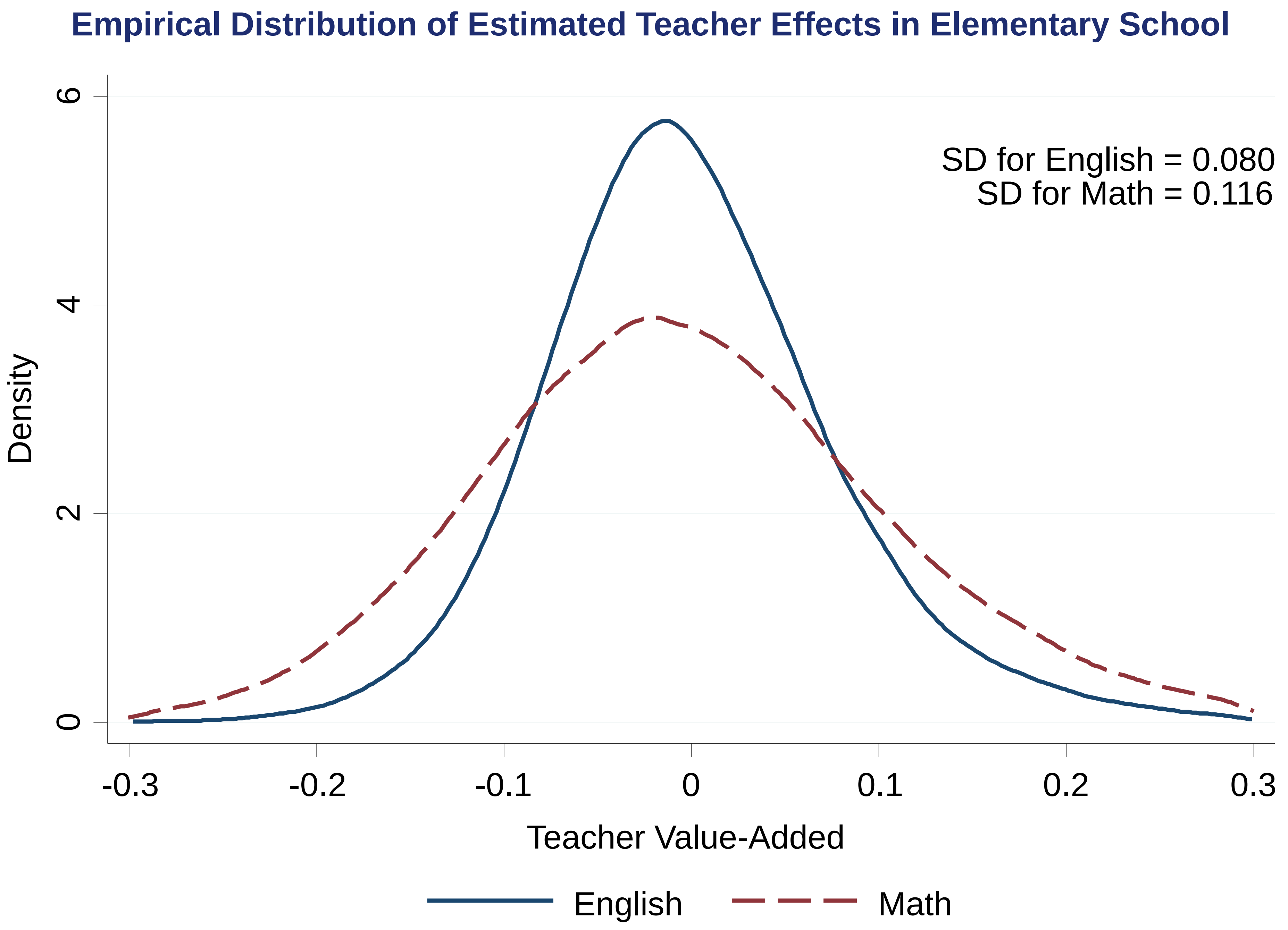
$$\hat{\mu}_{jt} = r_s \bar{A}_{j,t-s} \text{ where } r_s = Corr(\bar{A}_t, \bar{A}_{t-s})$$

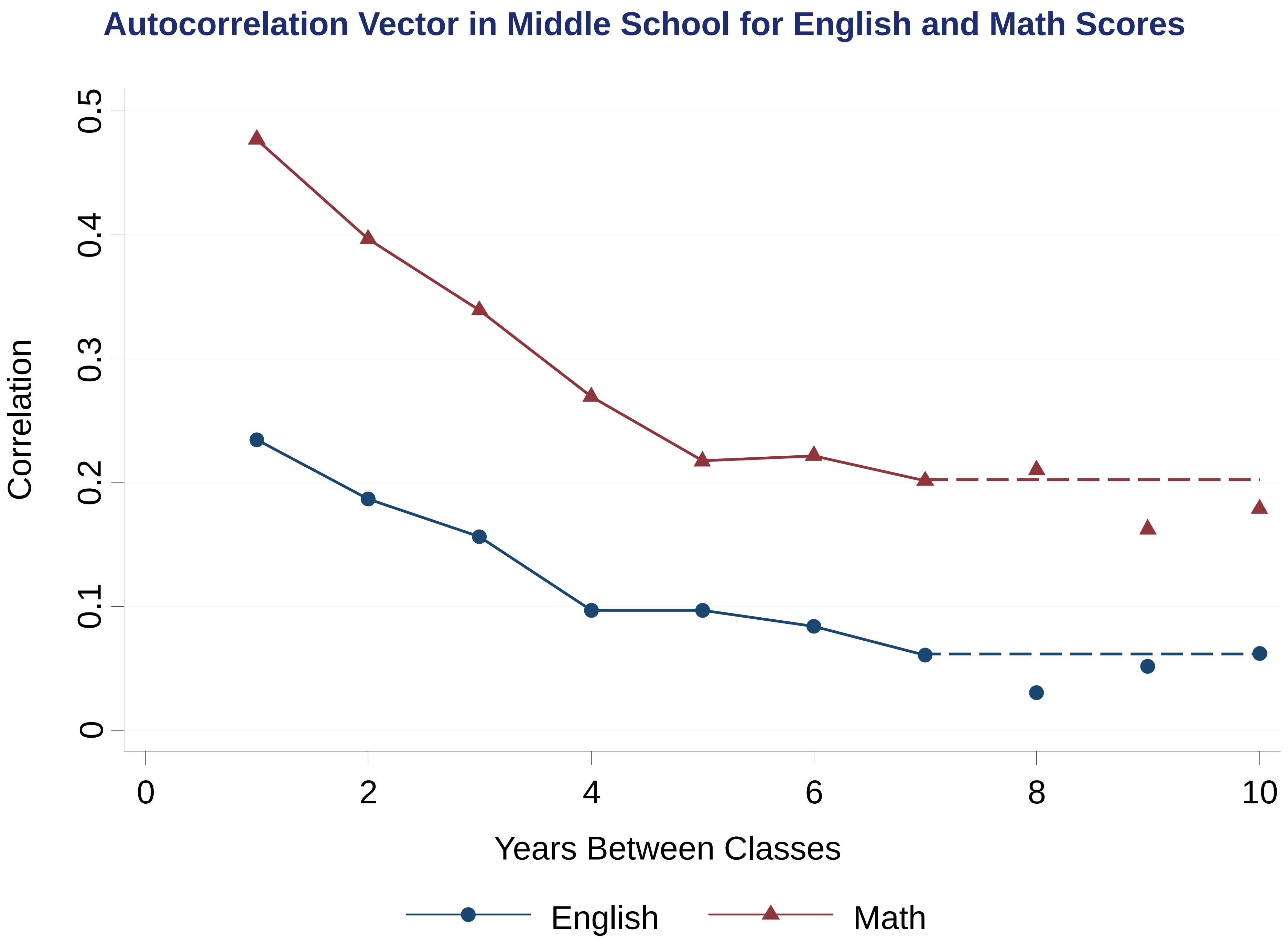
2. Without drift, put equal weight on all prior scores. Formula collapses to standard shrinkage estimator [e.g., Kane and Staiger 2008]

$$\hat{\mu}_{jt} = \bar{A}_j^{-t} \frac{\sigma_\mu^2}{\sigma_\mu^2 + (\sigma_\theta^2 + \sigma_{\tilde{\varepsilon}}^2/n)/T}$$

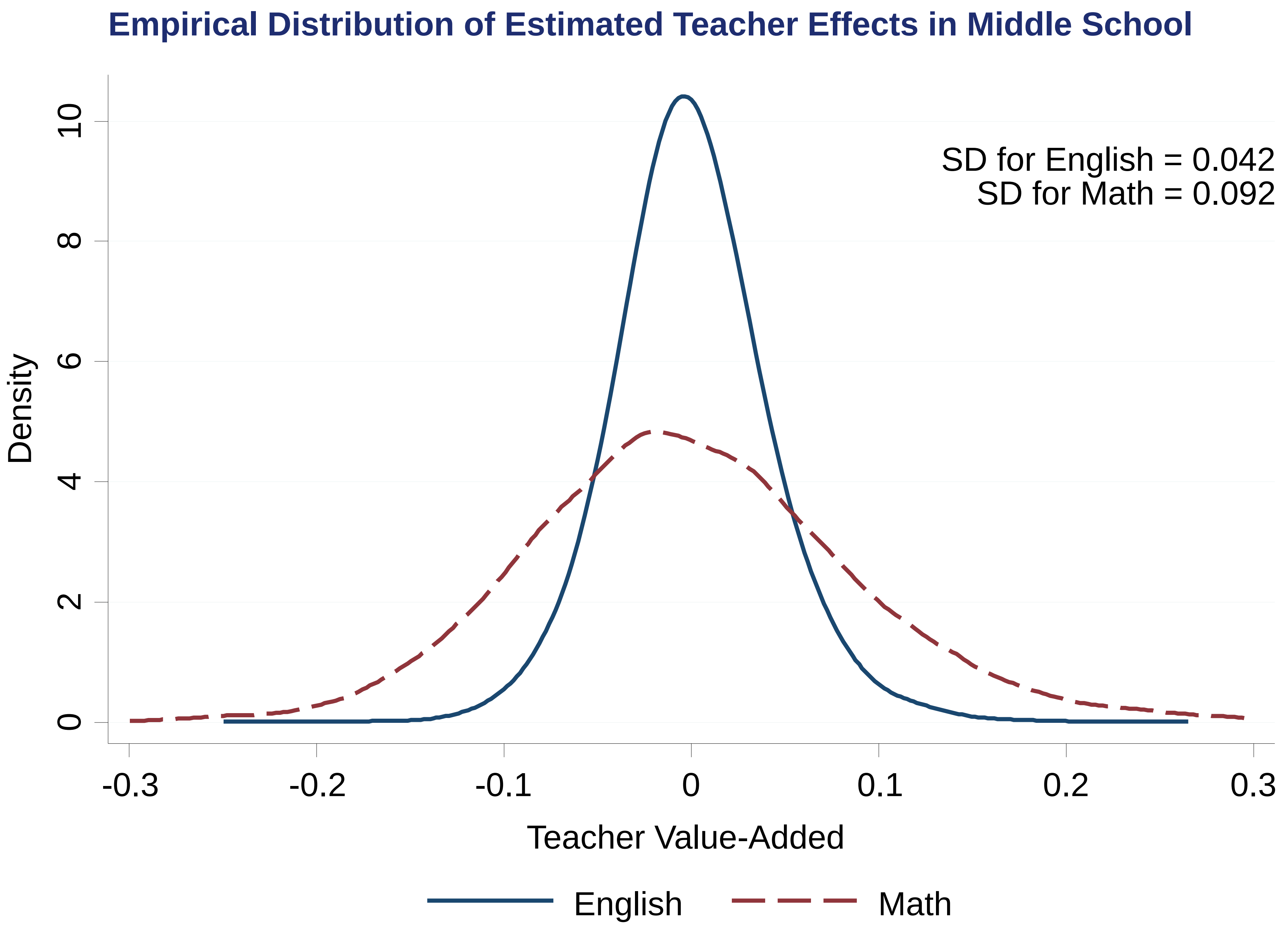


Autocorrelation Vector in Elementary School for English and Math Scores
Correlation (r_s)
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
Years Between Classes
English
Math

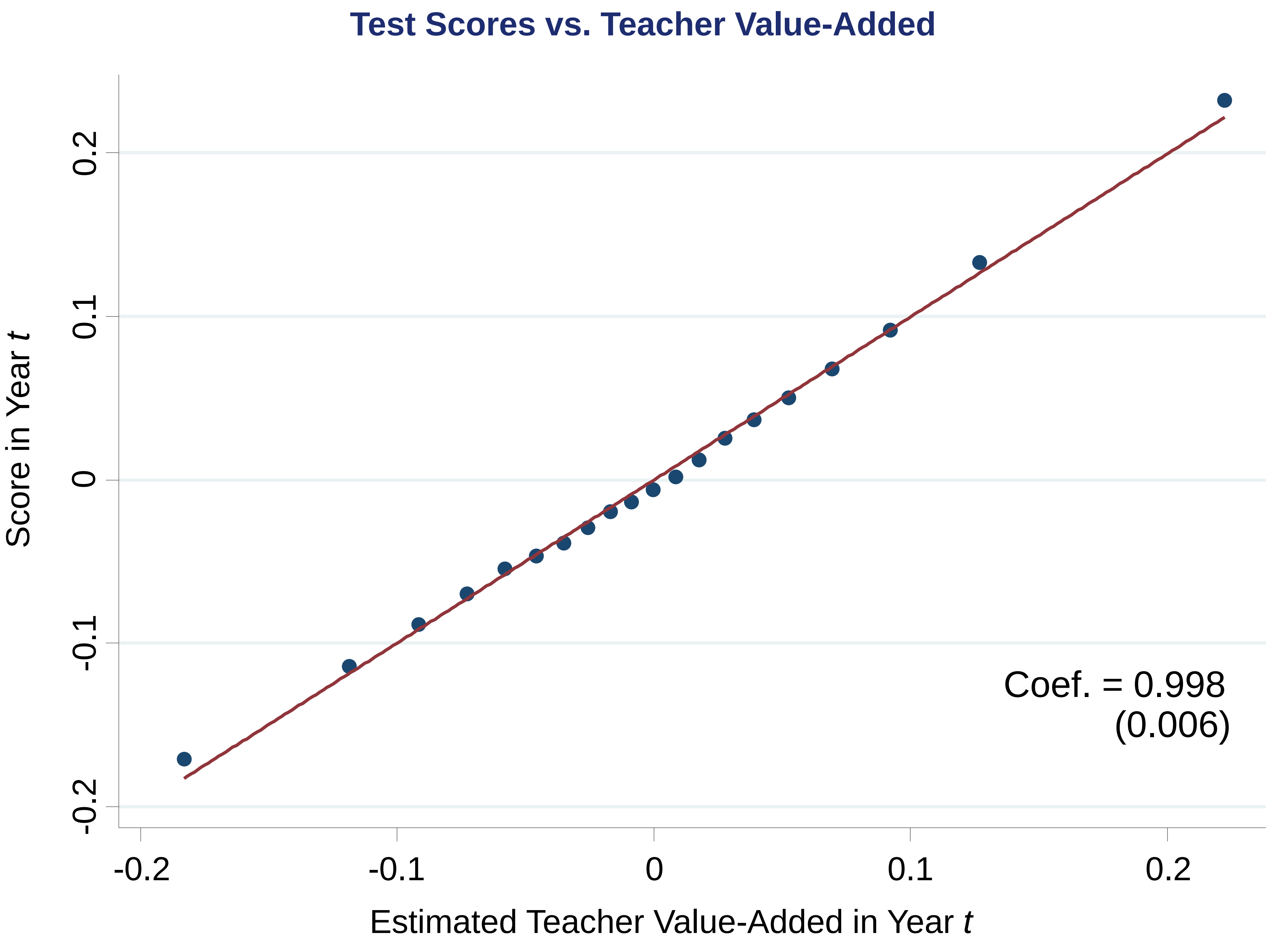

Empirical Distribution of Estimated Teacher Effects in Elementary School
SD for English = 0.080
SD for Math = 0.116
Density
6
4
2
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Value-Added
English
Math



Autocorrelation Vector in Middle School for English and Math Scores
Correlation
0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
Years Between Classes
English
Math



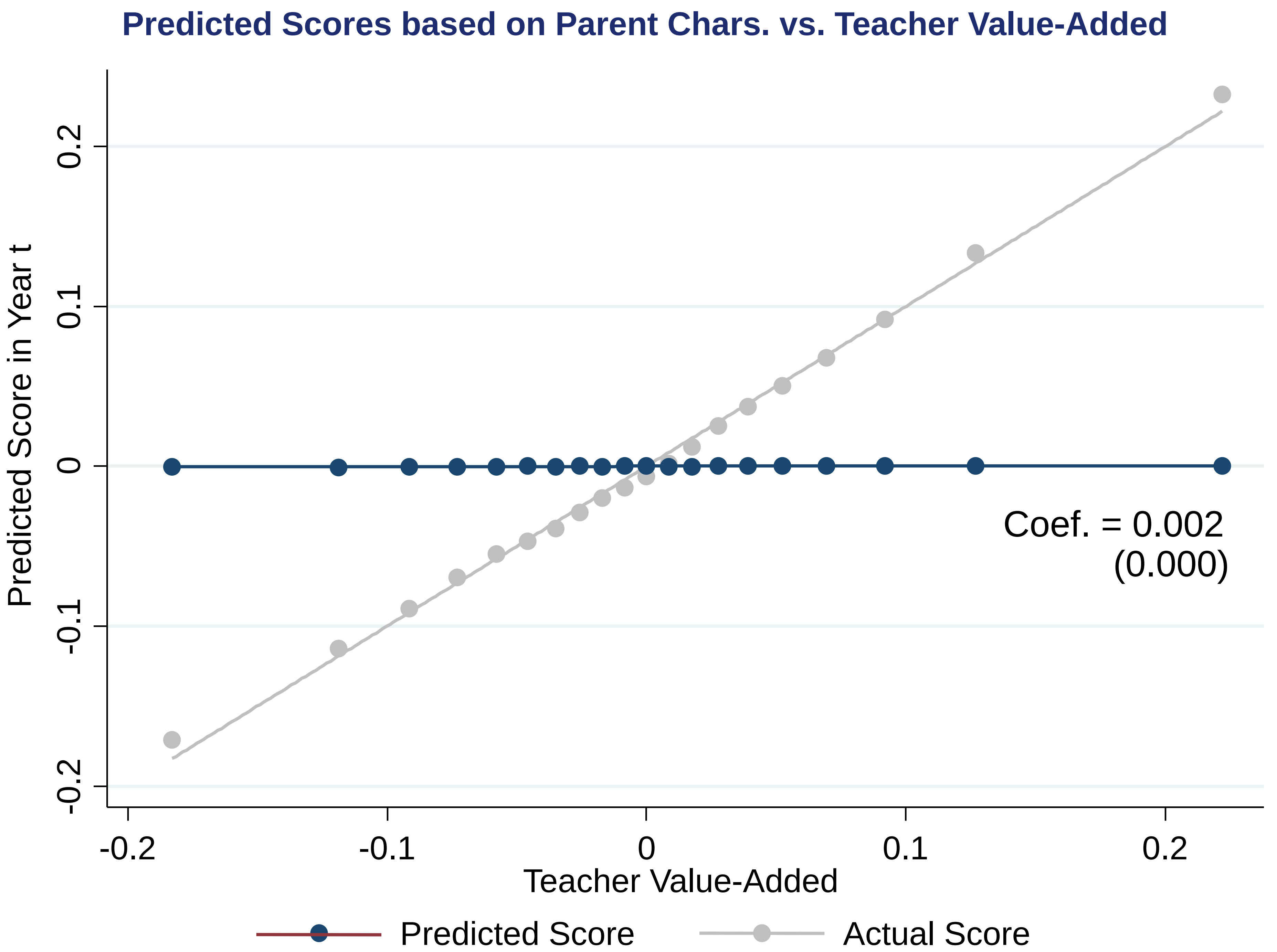
Empirical Distribution of Estimated Teacher Effects in Middle School
SD for English = 0.042
SD for Math = 0.092
Density
10
8
6
4
2
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Value-Added
English
Math


Test Scores vs. Teacher Value-Added
Score in Year t
0.2
0.1
0
-0.1
-0.2
-0.2
-0.1
0
0.1
0.2
Estimated Teacher Value-Added in Year t
Coef. = 0.998
(0.006)

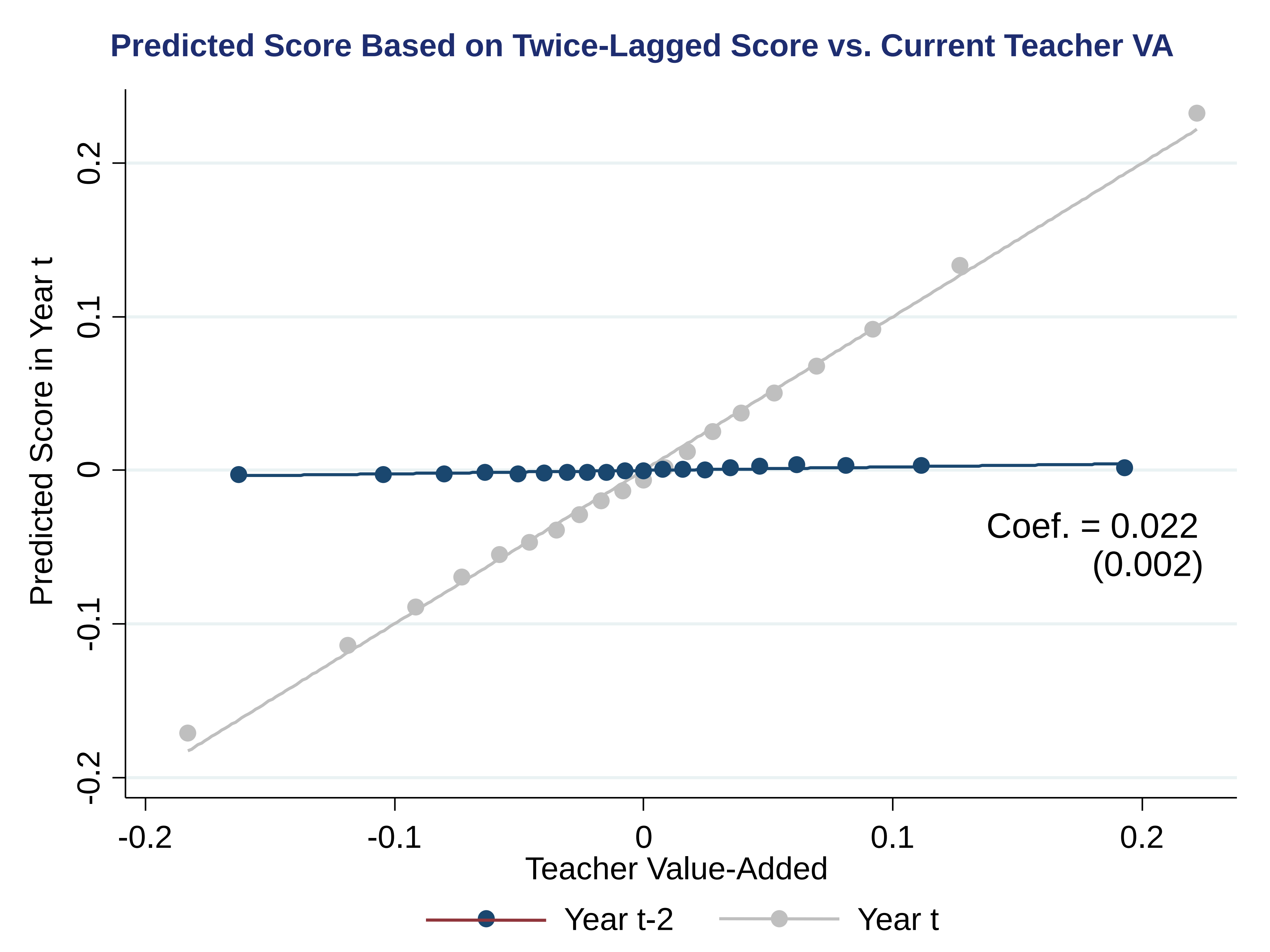
Part I: Bias in VA Estimates

Question 1: Are VA Estimates Unbiased?

- Teachers' estimated VA may reflect unobserved differences in type of students they get rather than causal impact of teacher
- We evaluate whether VA measures provided unbiased forecasts of teachers' causal impacts in two ways
- First test: are observable characteristics excluded from VA model are correlated with VA estimates?
    - Ex: parent income is a strong predictor of test scores even conditional on control vector used to estimate VA
    - Do high VA teachers have students from higher-income families?
    - Combine parental background characteristics into a single **predicted score** using a cross-sectional 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
Coef. = 0.002
(0.000)
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added
Predicted Score
Actual Score


Predicted Score Based on Twice-Lagged Score vs. Current Teacher VA
Predicted Score in Year t
0.2
0.1
0
-0.1
-0.2
Coef. = 0.022
(0.002)
-0.2
-0.1
0
0.1
0.2
Teacher Value-Added
Year t-2
Year t

Estimates of Forecast Bias Using Parent Characteristics and Lagged Scores

Dep. Var.:	Score in Year t	Pred. Score using Parent Chars.	Score in Year t	Pred. Score using Year t-2 Score
	(1)	(2)	(3)	(4)
Teacher VA	0.998	0.002	0.996	0.022
	(0.0057)	(0.0003)	(0.0057)	(0.0019)
Parent Chars. Controls			X	
Observations	6,942,979	6,942,979	6,942,979	5,096,518

Quasi-Experimental Validation: Teacher Switchers

- VA measures orthogonal to predictors of scores such as parent income

- But selection on unobservables could still be a problem (Rothstein 2010)

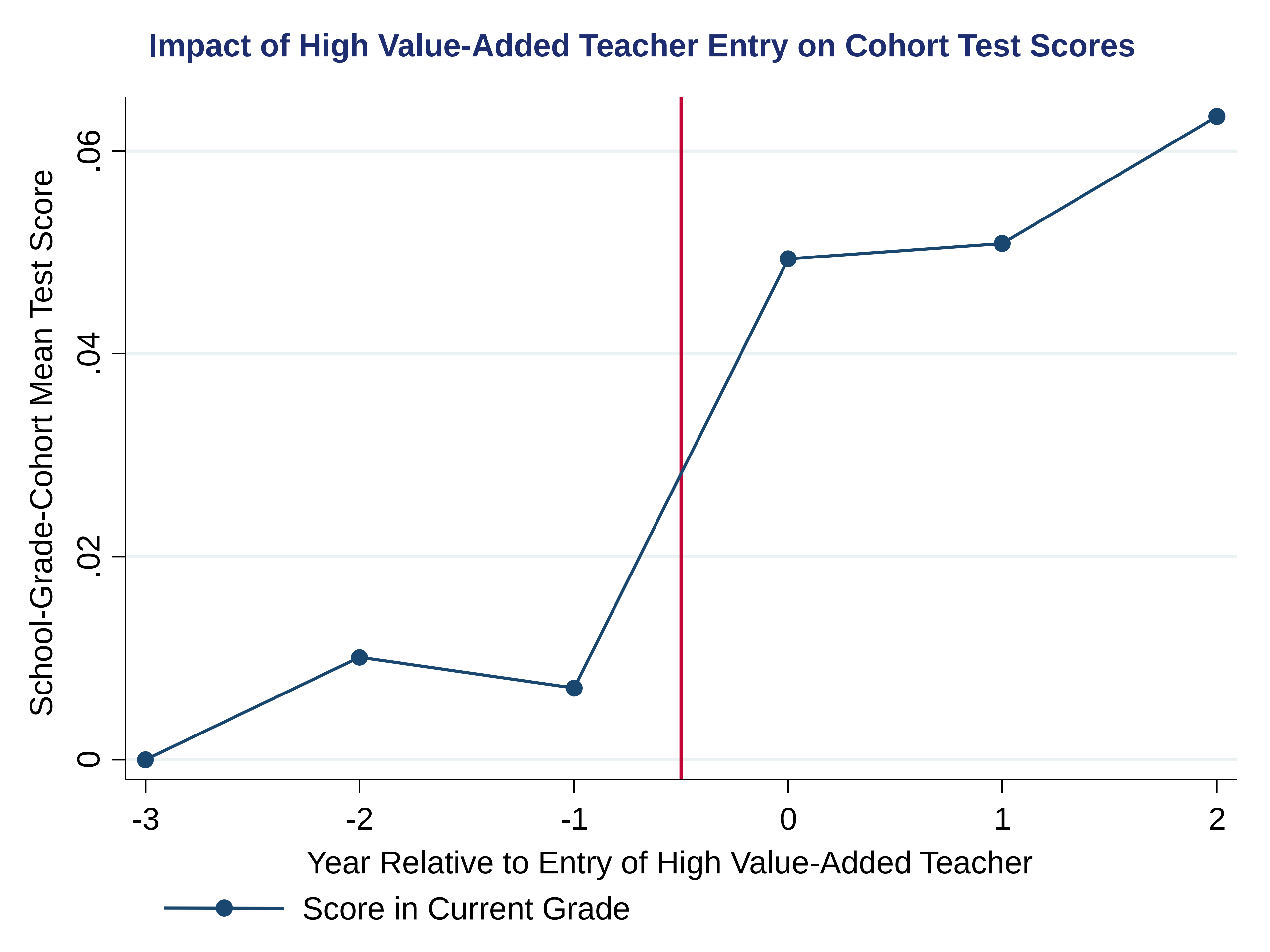
- Ideal test: out-of-sample forecasts in experiments (Kane and Staiger 2008)
  - Does a student who is randomly assigned to a teacher previously estimated to be high VA have higher test score gains?

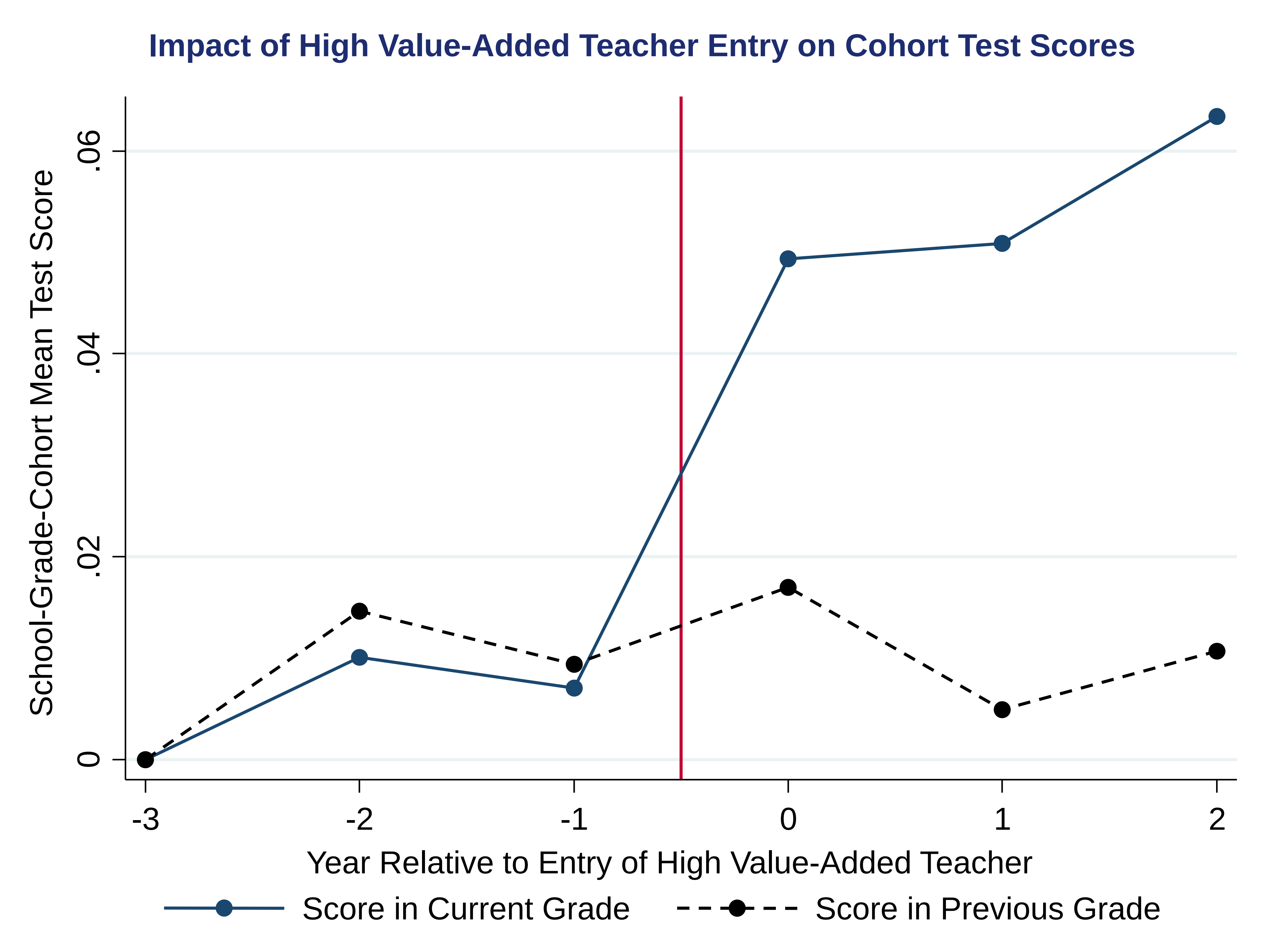
- We use teacher switching as a quasi-experimental analog

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

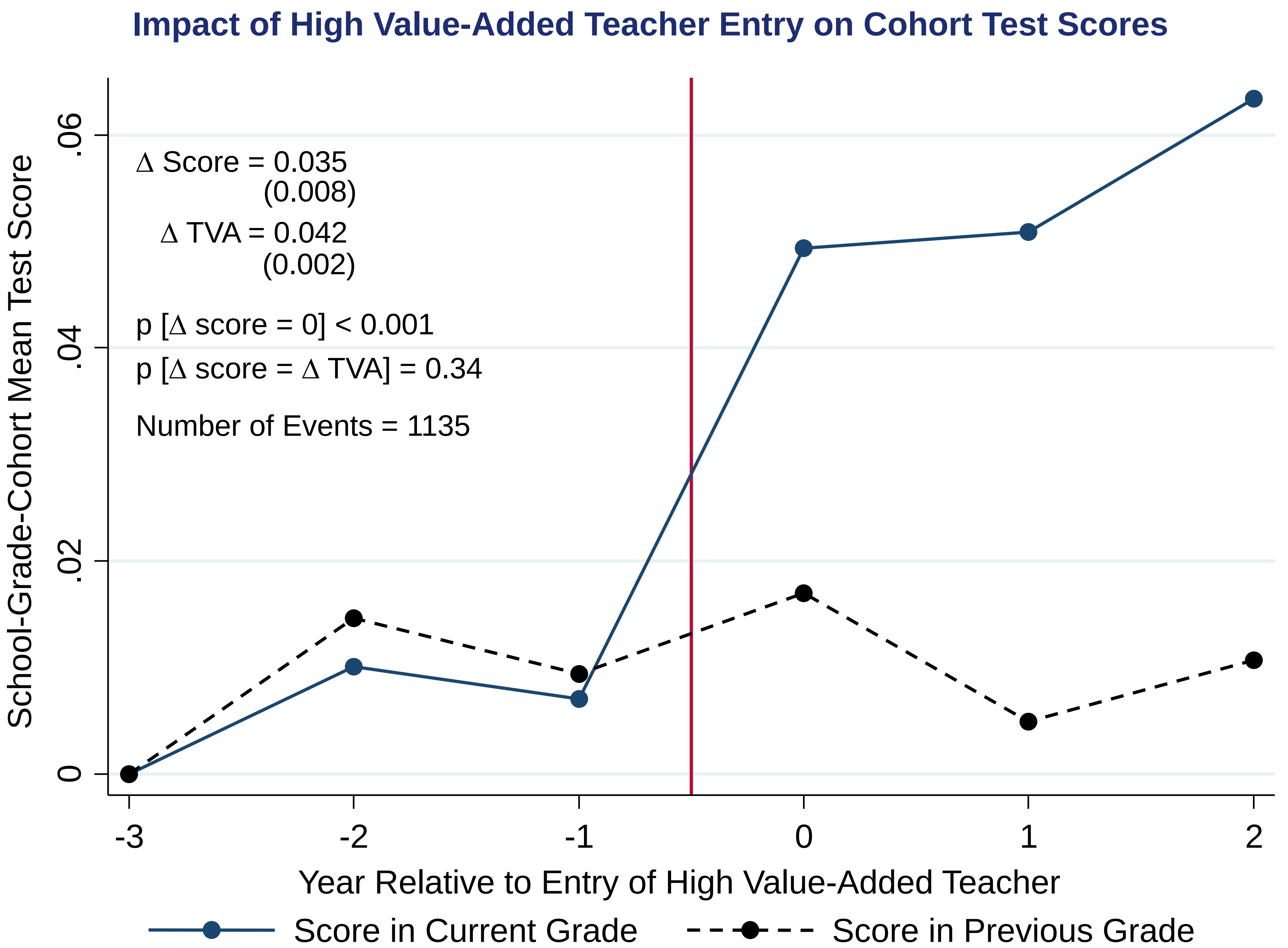
School	Grade	Subject	Year	Teachers	Mean Score	Mean Age 28 Earnings
1	5	math	1992	Smith, Farber, …	-.09	$15K
1	5	math	1993	Smith, Farber, …	-.04	$17K
1	5	math	1994	Smith, Farber, …	-.05	$16K
1	5	math	1995	Mas, Farber, …	0.01	$18K
1	5	math	1996	Mas, Farber, …	0.04	$17K
1	5	math	1997	Mas, Farber, …	0.02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

- Smith switches to a different school in 1995; Mas replaces him

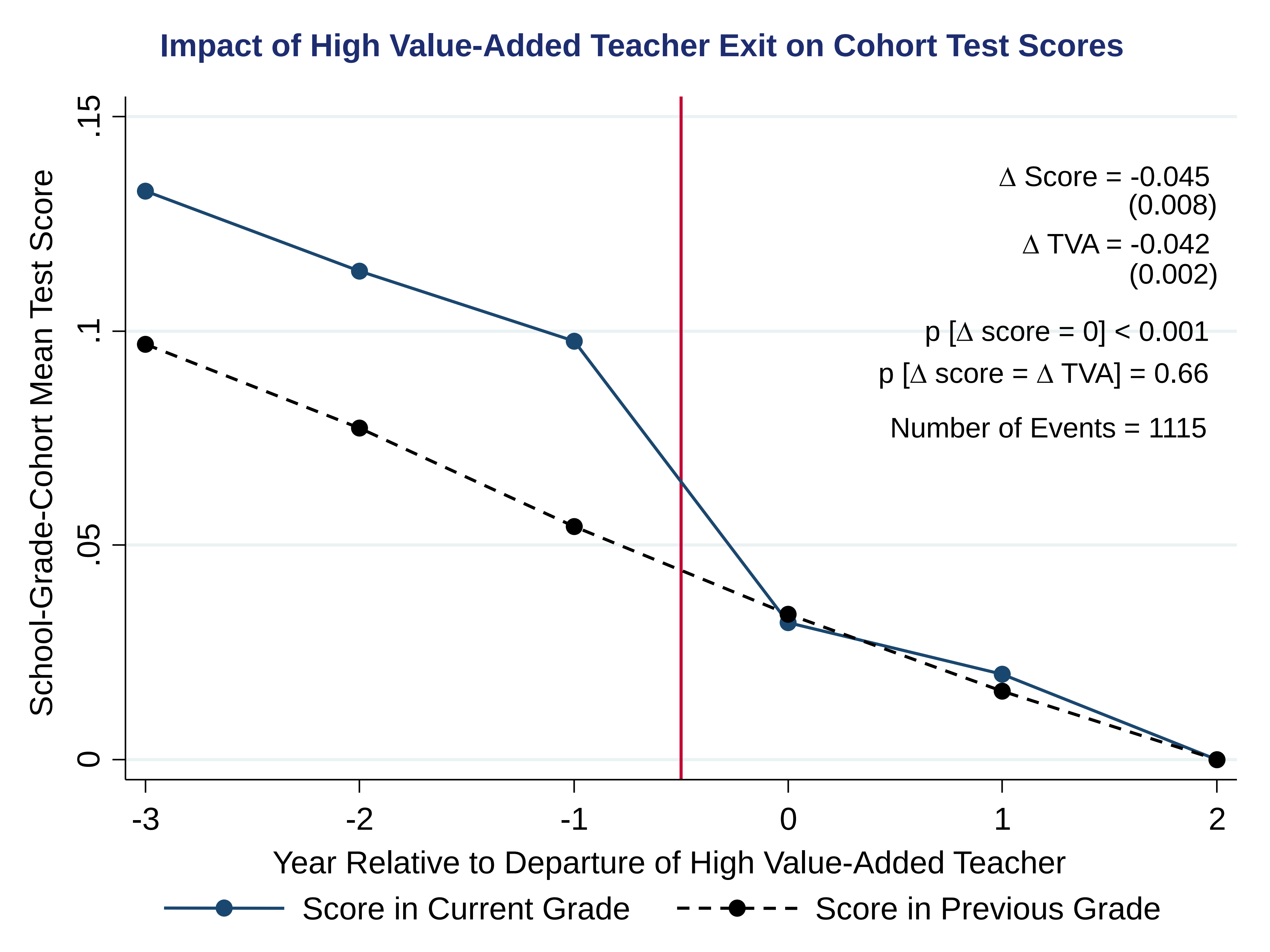

Impact of High Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0
.02
.04
.06
-3
-2
-1
0
1
2
Year Relative to Entry of High Value-Added Teacher
Score in Current Grade

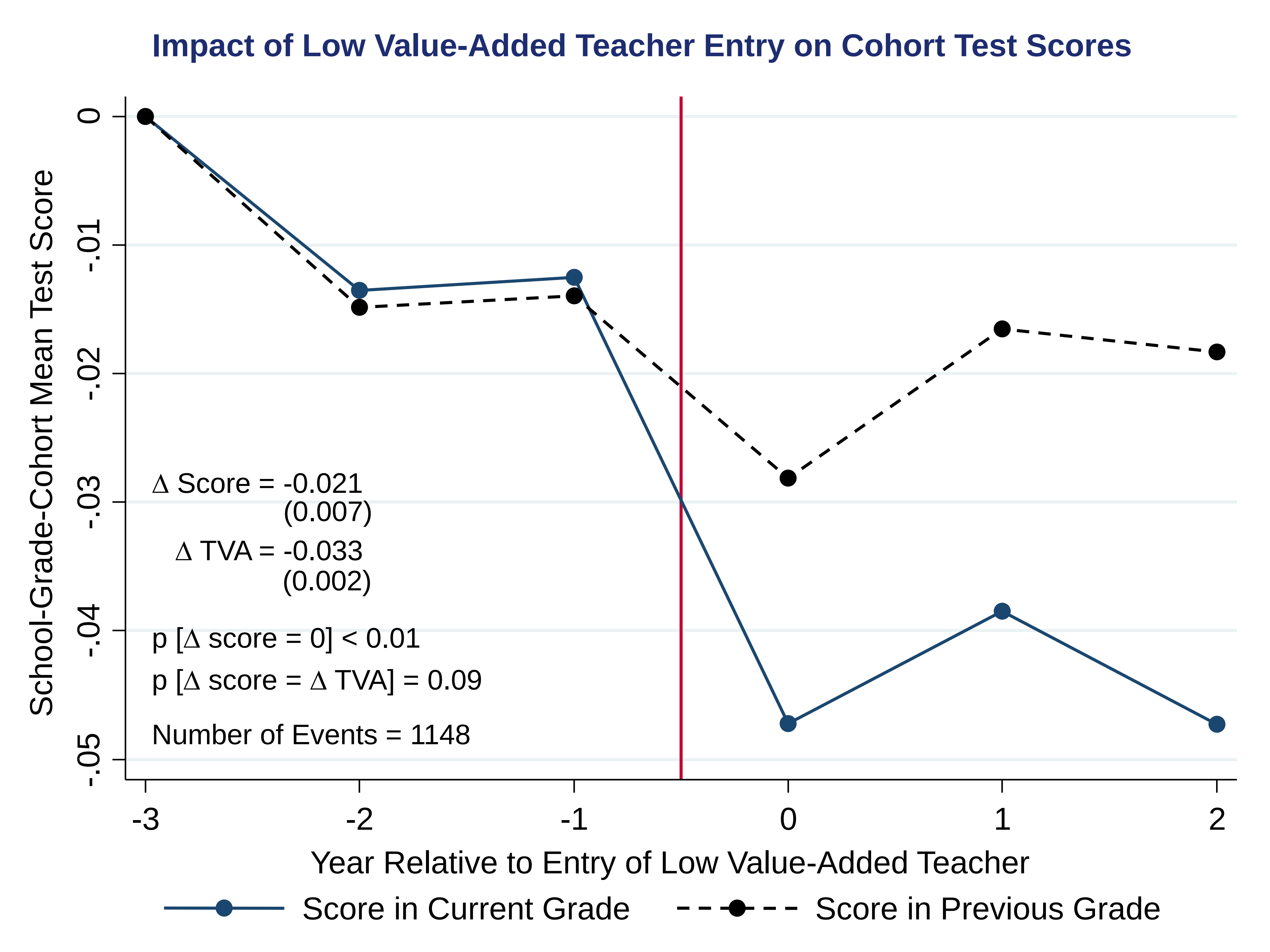


Impact of High Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0
.02
.04
.06
-3
-2
-1
0
1
2
Year Relative to Entry of High Value-Added Teacher
Score in Current Grade
Score in Previous Grade

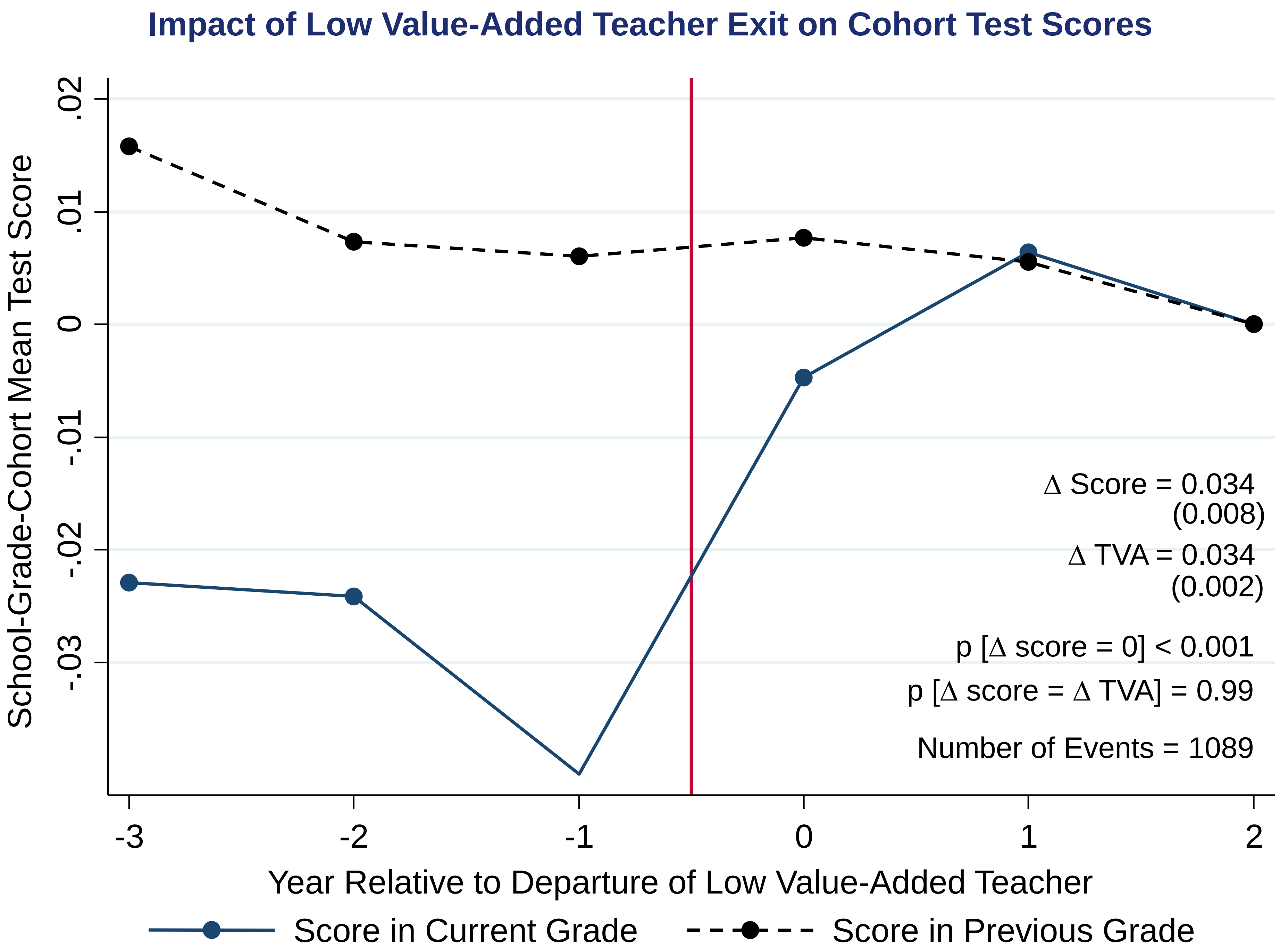


Impact of High Value-Added Teacher Entry on Cohort Test Scores
Δ Score = 0.035
(0.008)
Δ TVA = 0.042
(0.002)
p [Δ score = 0] < 0.001
p [Δ score = Δ TVA] = 0.34
Number of Events = 1135
School-Grade-Cohort Mean Test Score
0
.02
.04
.06
-3
-2
-1
0
1
2
Year Relative to Entry of High Value-Added Teacher
Score in Current Grade
Score in Previous Grade

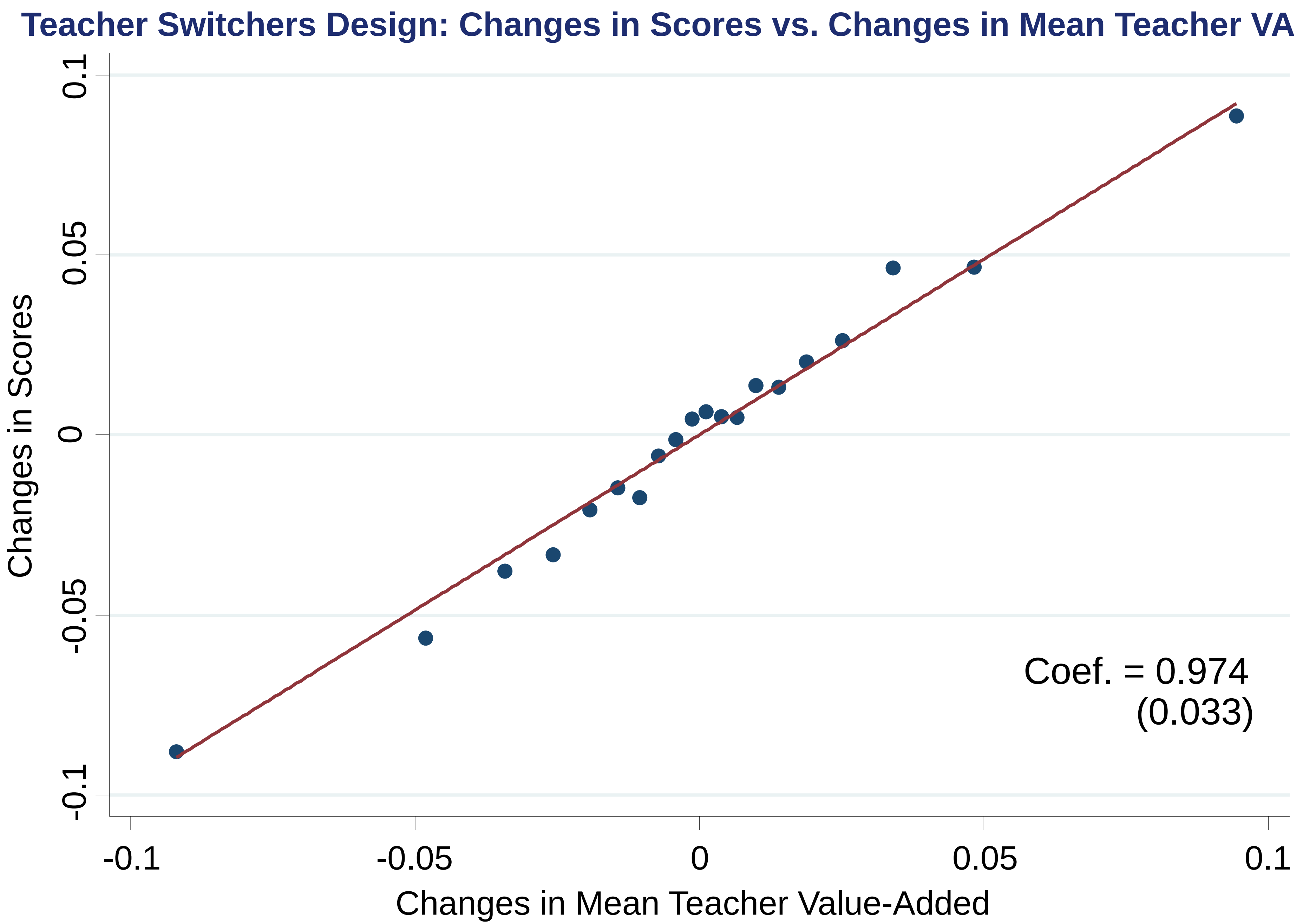

Impact of High Value-Added Teacher Exit on Cohort Test Scores
Δ Score = -0.045
(0.008)
Δ TVA = -0.042
(0.002)
p [Δ score = 0] < 0.001
p [Δ score = Δ TVA] = 0.66
Number of Events = 1115
School-Grade-Cohort Mean Test Score
0
.05
.1
.15
-3
-2
-1
0
1
2
Year Relative to Departure of High Value-Added Teacher
Score in Current Grade
Score in Previous Grade



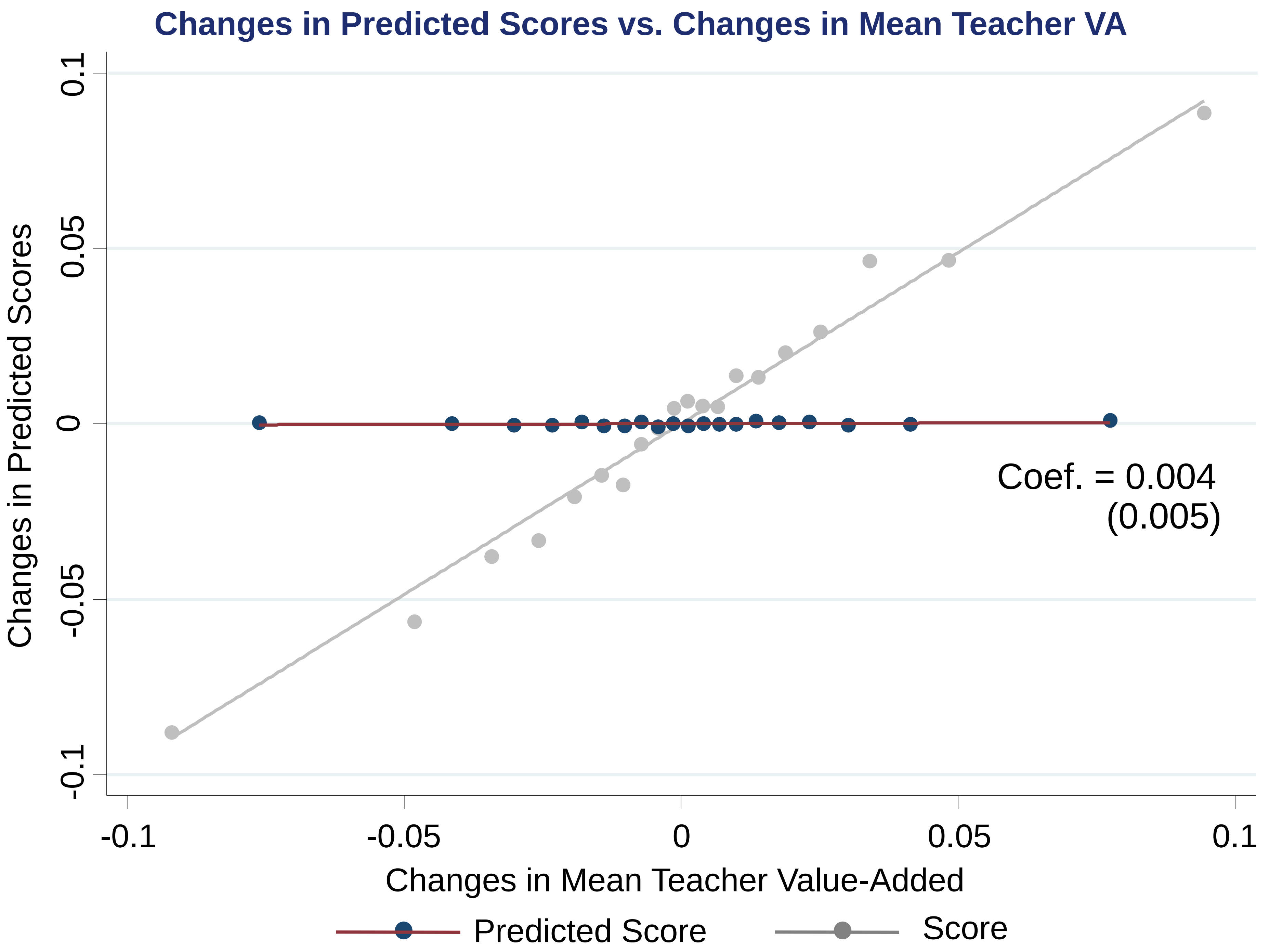
Impact of Low Value-Added Teacher Entry on Cohort Test Scores
School-Grade-Cohort Mean Test Score
0
-.01
-.02
-.03
-.04
-.05
Δ Score = -0.021
(0.007)
Δ TVA = -0.033
(0.002)
p [Δ score = 0] < 0.01
p [Δ score = Δ TVA] = 0.09
Number of Events = 1148
-3
-2
-1
0
1
2
Year Relative to Entry of Low Value-Added Teacher
Score in Current Grade
Score in Previous Grade

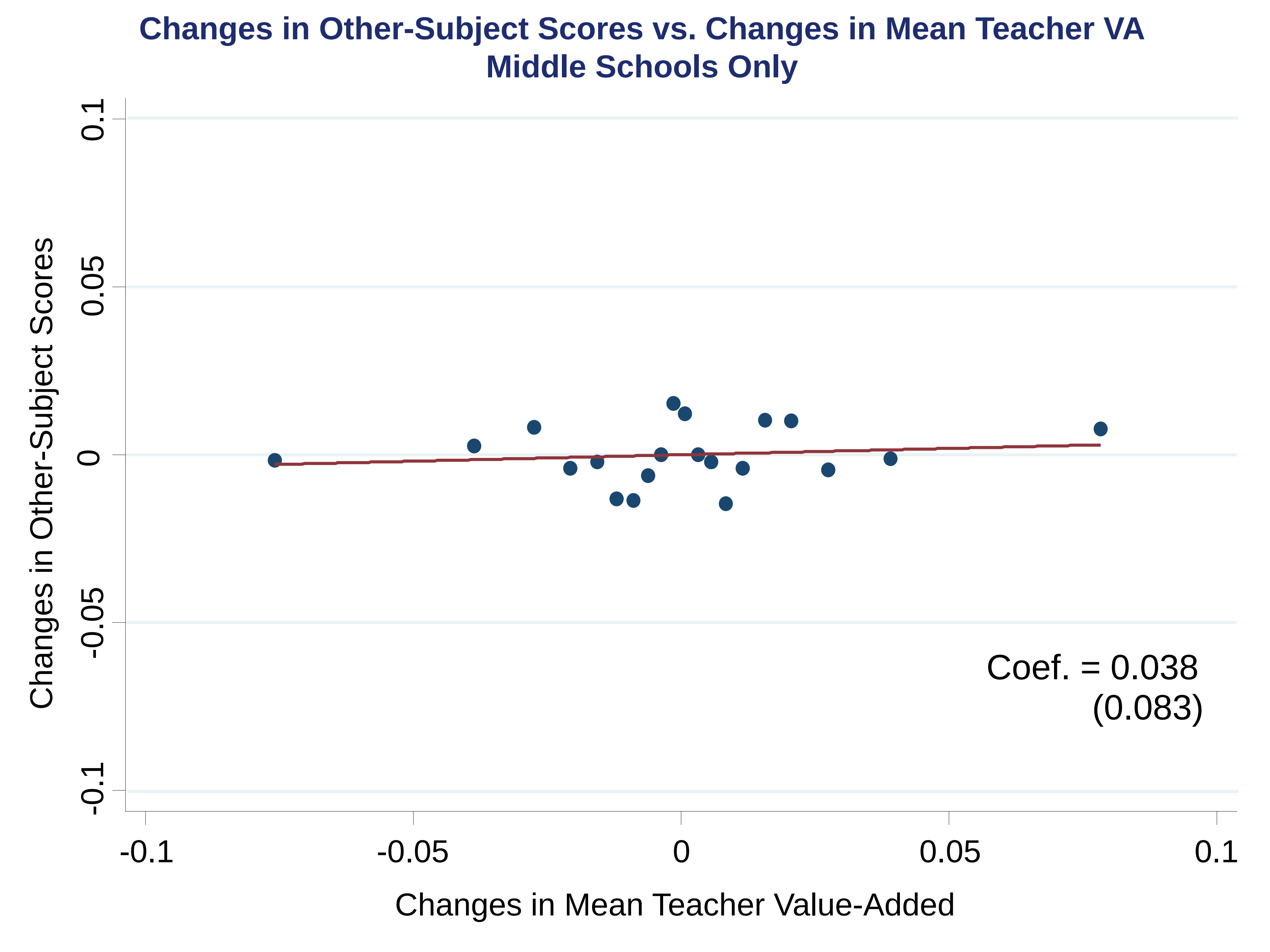


Impact of Low Value-Added Teacher Exit on Cohort Test Scores
School-Grade-Cohort Mean Test Score
.02
.01
0
-.01
-.02
-.03
-3
-2
-1
0
1
2
Year Relative to Departure of Low Value-Added Teacher
Δ Score = 0.034
(0.008)
Δ TVA = 0.034
(0.002)
p [Δ score = 0] < 0.001
p [Δ score = Δ TVA] = 0.99
Number of Events = 1089
Score in Current Grade
Score in Previous Grade

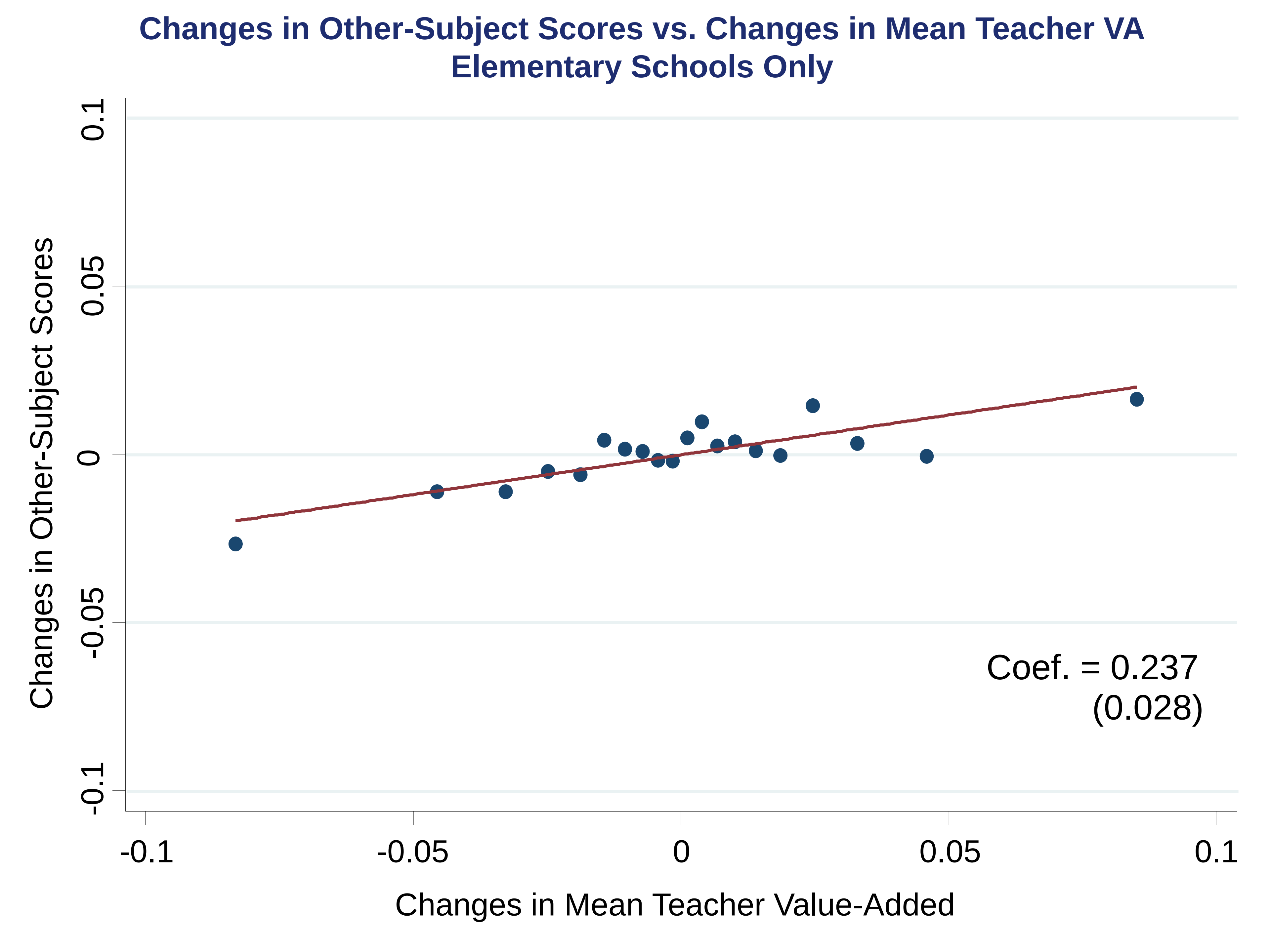


Teacher Switchers Design: Changes in Scores vs. Changes in Mean Teacher VA
Changes in Scores
0.1
0.05
0
-0.05
-0.1
-0.1
-0.05
0
0.05
0.1
Changes in Mean Teacher Value-Added
Coef. = 0.974
(0.033)


Changes in Predicted Scores vs. Changes in Mean Teacher VA
Changes in Predicted Scores
0.1
0.05
0
-0.05
-0.1
-0.1
-0.05
0
0.05
0.1
Changes in Mean Teacher Value-Added
Coef. = 0.004
(0.005)
Predicted Score
Score



Changes in Other-Subject Scores vs. Changes in Mean Teacher VA
Middle Schools Only
Changes in Other-Subject Scores
0.1
0.05
0
-0.05
-0.1
Coef. = 0.038
(0.083)
-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
Elementary Schools Only
Changes in Other-Subject Scores
0.1
0.05
0
-0.05
-0.1
-0.1
-0.05
0
0.05
0.1
Changes in Mean Teacher Value-Added
Coef. = 0.237
(0.028)

Estimates of Forecast Bias with Alternative Control Vectors

Control Vector	Quasi-Experimental Estimate of Bias (%)
Baseline	2.58 (3.34)
Student-level lagged scores	4.83 (3.29)
Non-score controls only	45.39 (2.26)
No controls	65.58 (3.73)

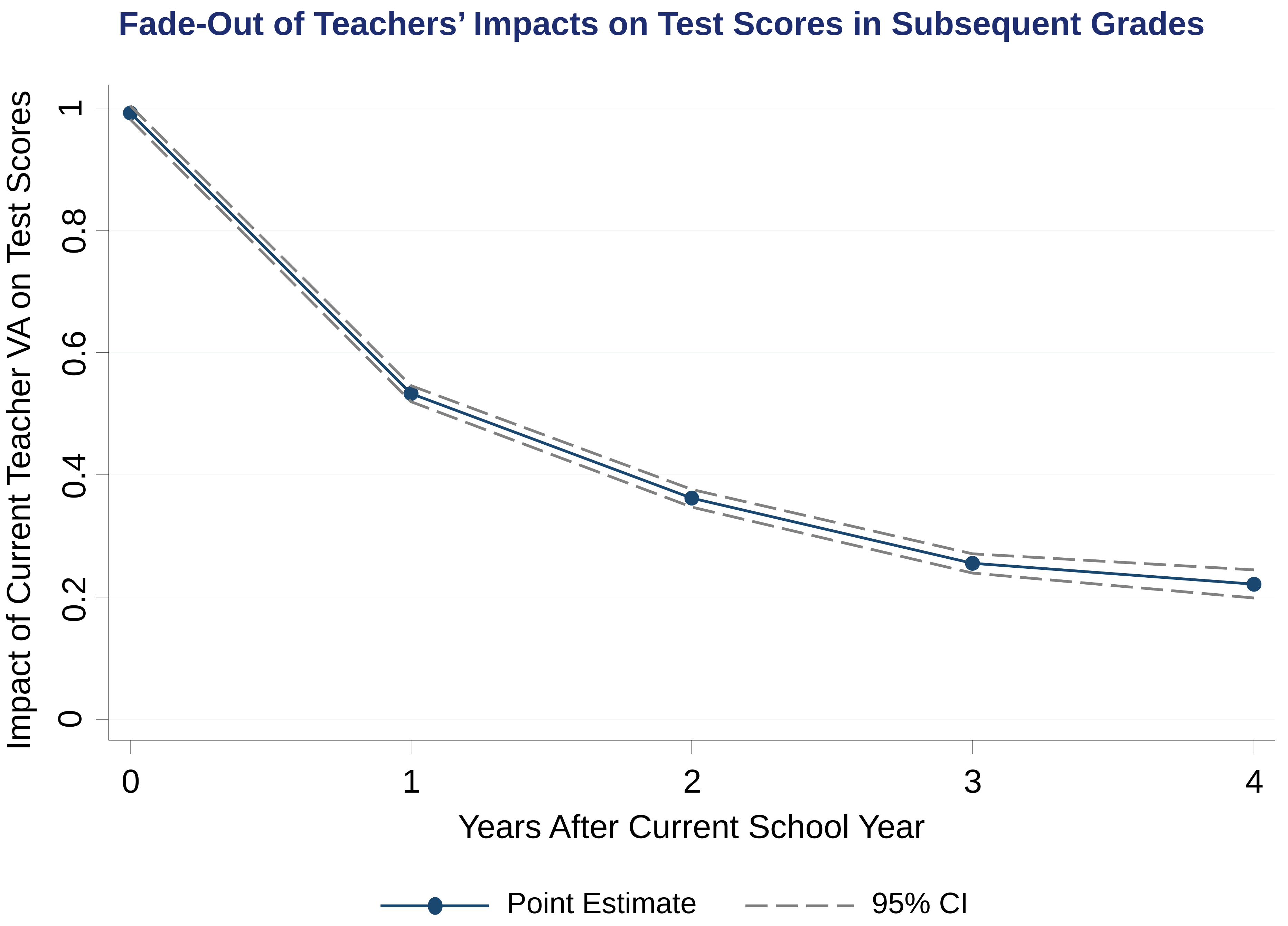
Relation to Rothstein (2010) Findings on Sorting

- Rothstein result 1: Students are sorted into classrooms based on pre-determined variables such as grade *g-2* test scores
  - We confirm this result in our data
- Rothstein result 2: Selection on observables is minimal conditional on grade *g-1* controls
  - Controlling for grade *g-2* score does not affect VA estimates
  - Consistent with our findings that VA does not predict *g-2* score

→ Rothstein notes that his findings do not imply bias in VA estimates

  - But they raise concerns about potential selection on unobservables
  - Our quasi-experimental teacher switcher tests indicate that selection on unobservables turns out to be modest in practice

Part II: Long-Term Impacts



Fade-Out of Teachers' Impacts on Test Scores in Subsequent Grades
Impact of Current Teacher VA on Test Scores
0
0.2
0.4
0.6
0.8
1
0
1
2
3
4
Years After Current School Year
Point Estimate
95% CI

Impacts on Outcomes in Adulthood

- Do teachers who raise test scores also improve long-term outcomes?

- Regress residualized long-term outcomes on teacher-level VA estimates

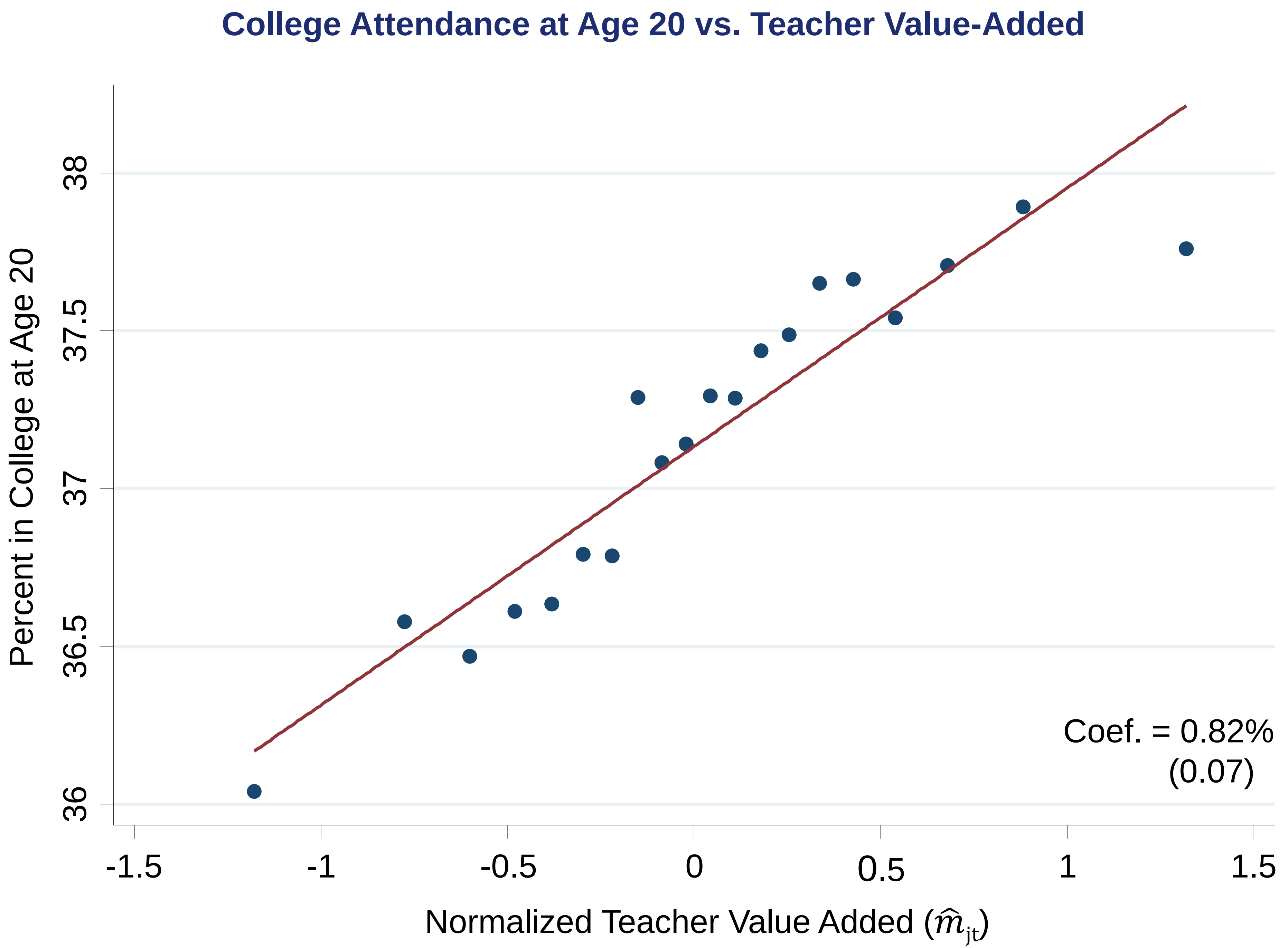
$$Y_{it} = \alpha + \kappa \hat{m}_{jt} + \eta'_{it}$$

- Then validate OLS estimates using cross-cohort switchers design

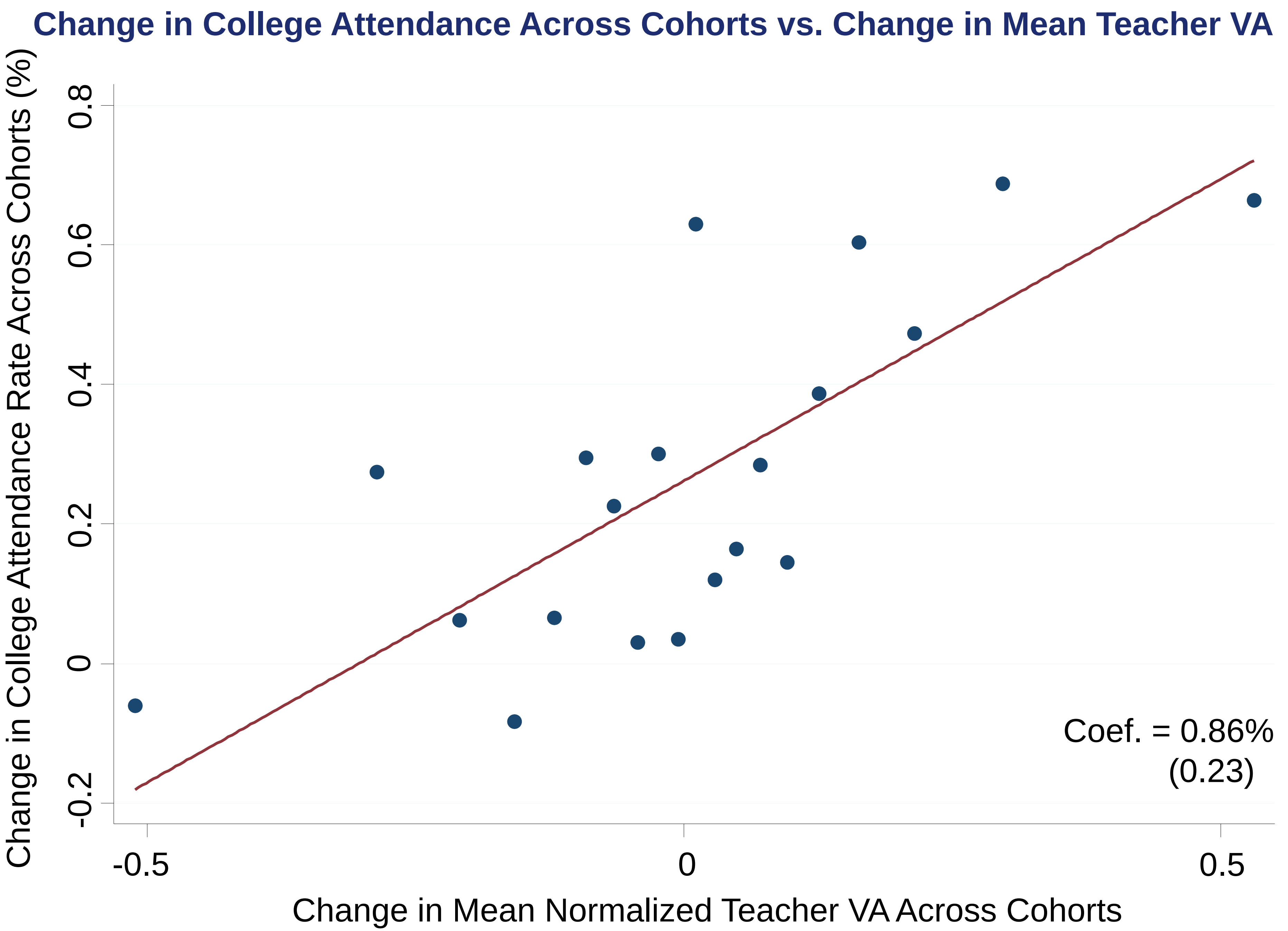
- Interpretation of these reduced-form coefficients [Todd and Wolpin 2003]

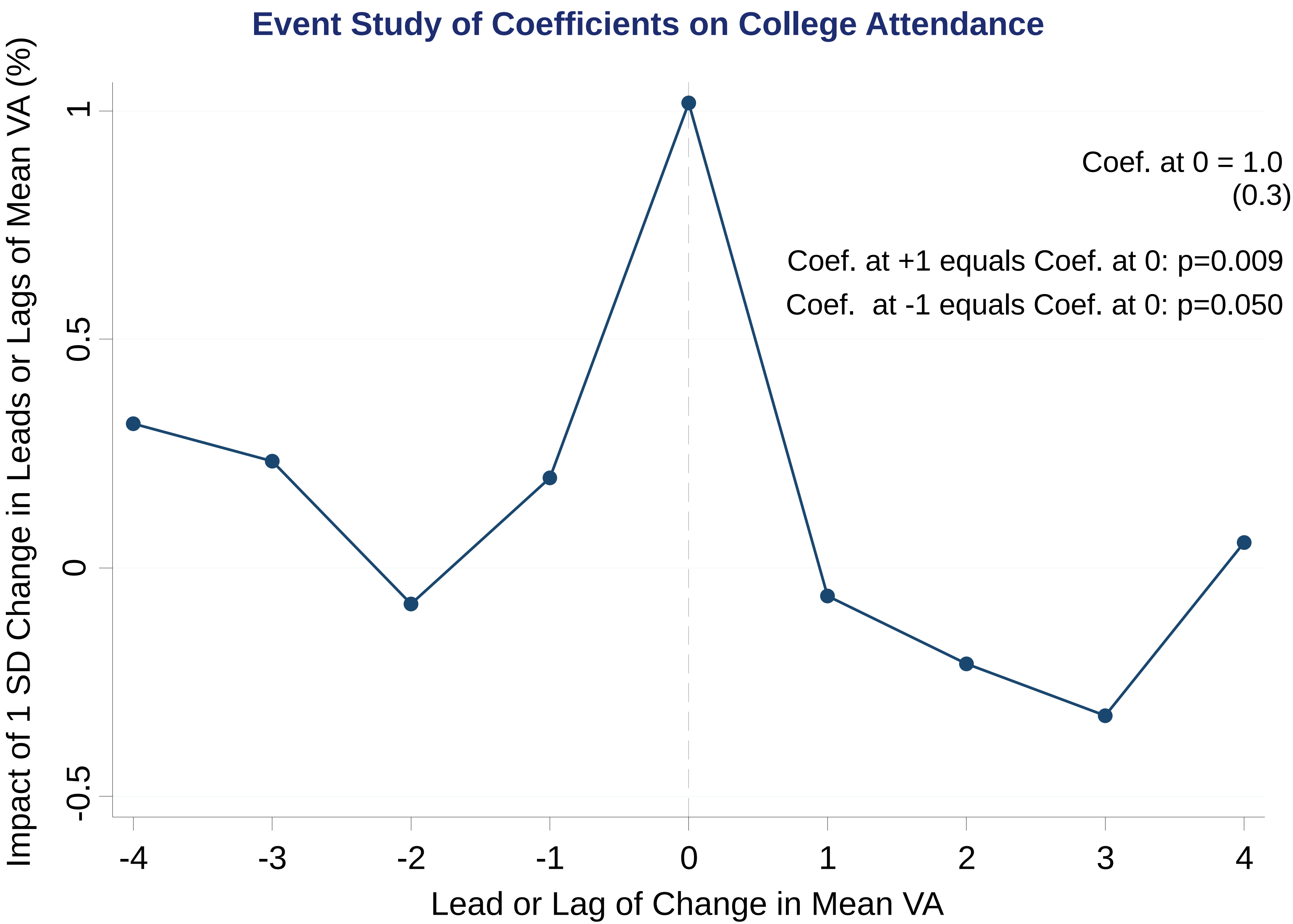
  - Impact of having better teacher, as measured by VA, for a **single year** during grades 4-8 on earnings

  - Includes benefit of better teachers, peers, etc. in later grades via tracking, as well as any complementarity with future teacher quality



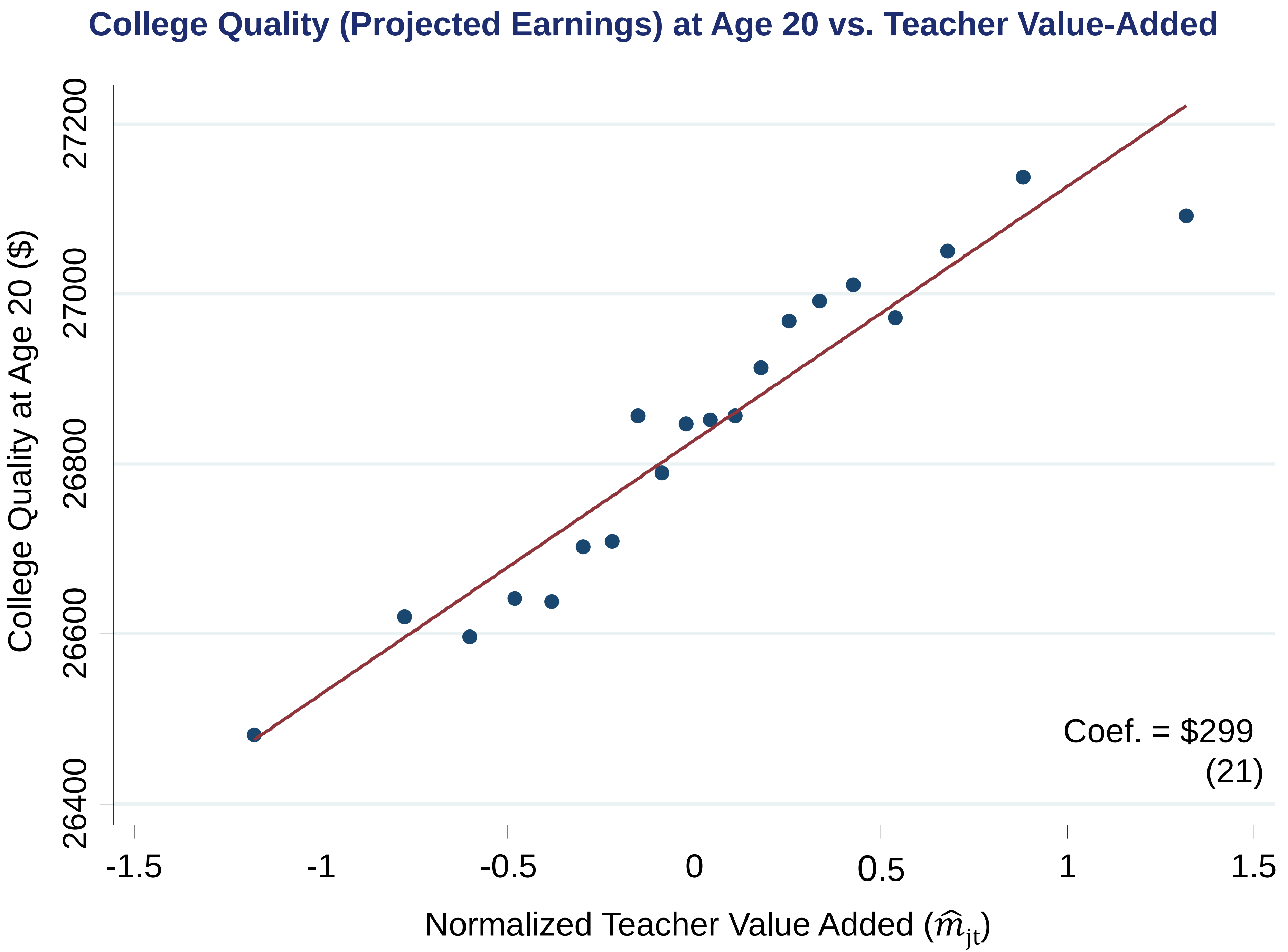
College Attendance at Age 20 vs. Teacher Value-Added
Percent in College at Age 20
38
37.5
37
36.5
36
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\hat{m}_{jt}$)
Coef. = 0.82%
(0.07)

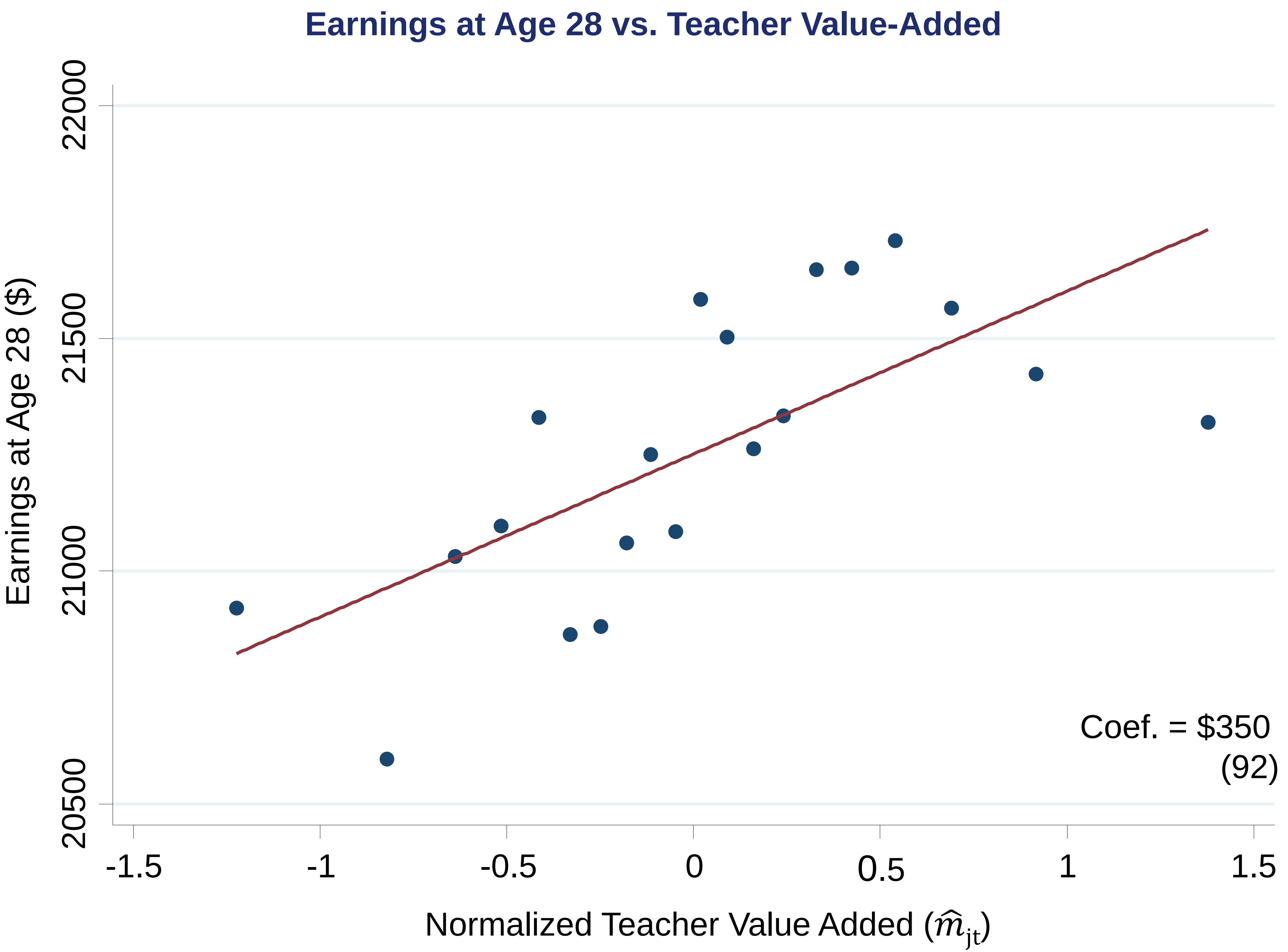

Change in College Attendance Across Cohorts vs. Change in Mean Teacher VA
Change in College Attendance Rate Across Cohorts (%)
0.8
0.6
0.4
0.2
0
-0.2
Coef. = 0.86%
(0.23)
-0.5
0
0.5
Change in Mean Normalized Teacher VA Across Cohorts


Event Study of Coefficients on College Attendance
Impact of 1 SD Change in Leads or Lags of Mean VA (%)
-0.5
0
0.5
1
-4
-3
-2
-1
0
1
2
3
4
Lead or Lag of Change in Mean VA
Coef. at 0 = 1.0
(0.3)
Coef. at +1 equals Coef. at 0: p=0.009
Coef. at -1 equals Coef. at 0: p=0.050

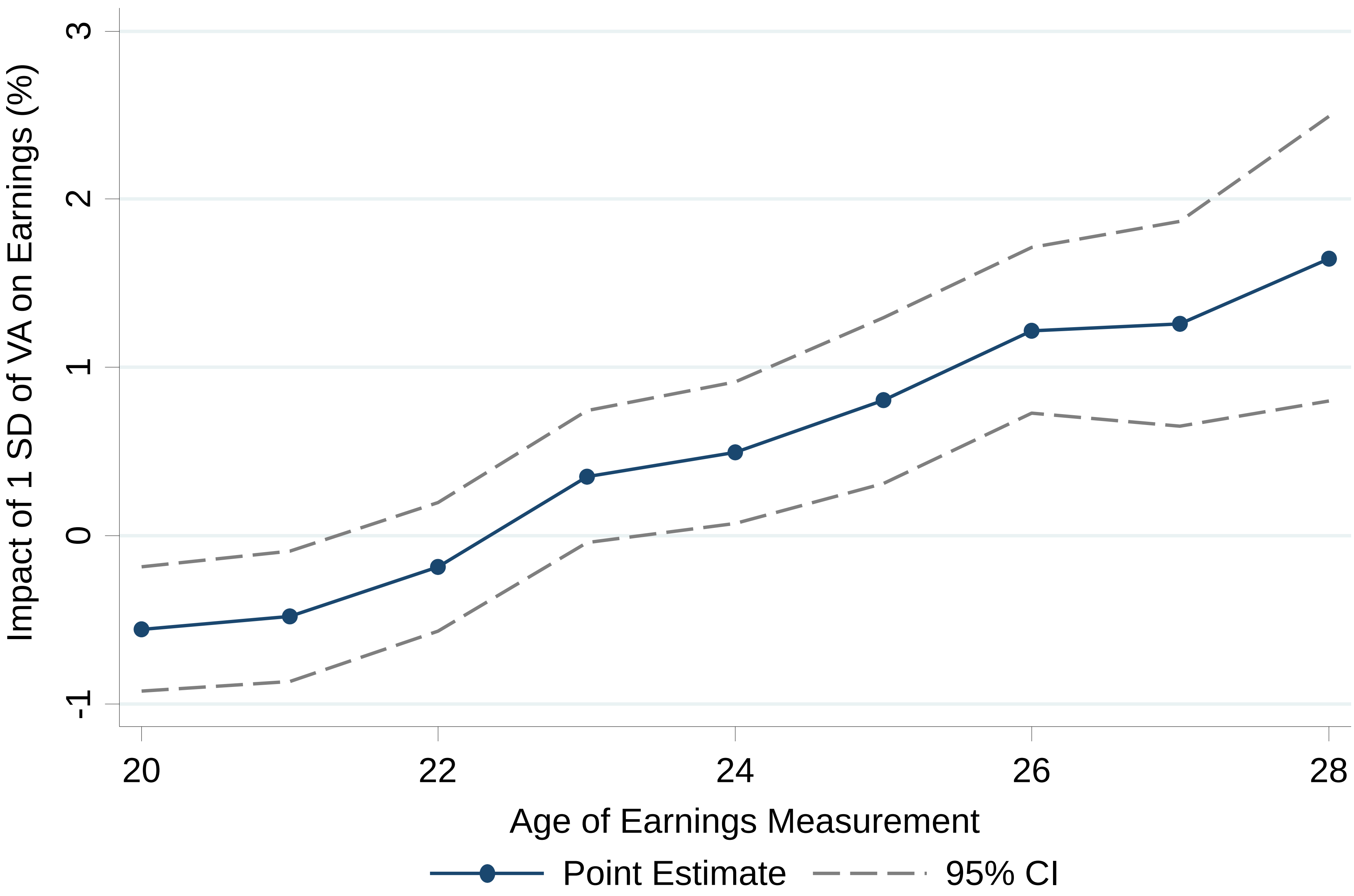
Impacts of Teacher Value-Added on College Attendance

Dependent Variable:	College at Age 20	College at Age 20	College at Age 20	College Quality at Age 20	College Quality at Age 20	College Quality at Age 20	High Quality College
	(%)	(%)	(%)	($)	($)	($)	(%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Value-Added	0.82 (0.07)	0.71 (0.06)	0.74 (0.09)	298.63 (20.74)	265.82 (18.31)	266.17 (26.03)	0.72 (0.05)
Mean of Dep. Var.	37.22	37.22	37.09	26,837	26,837	26,798	13.41
Baseline Controls	X	X	X	X	X	X	X
Parent Chars. Controls		X			X		
Lagged Score Controls			X			X	
Observations	4,170,905	4,170,905	3,130,855	4,167,571	4,167,571	3,128,478	4,167,571


College Quality (Projected Earnings) at Age 20 vs. Teacher Value-Added
College Quality at Age 20 ($)
27200
27000
26800
26600
26400
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\hat{m}_{jt}$)
Coef. = $299
(21)

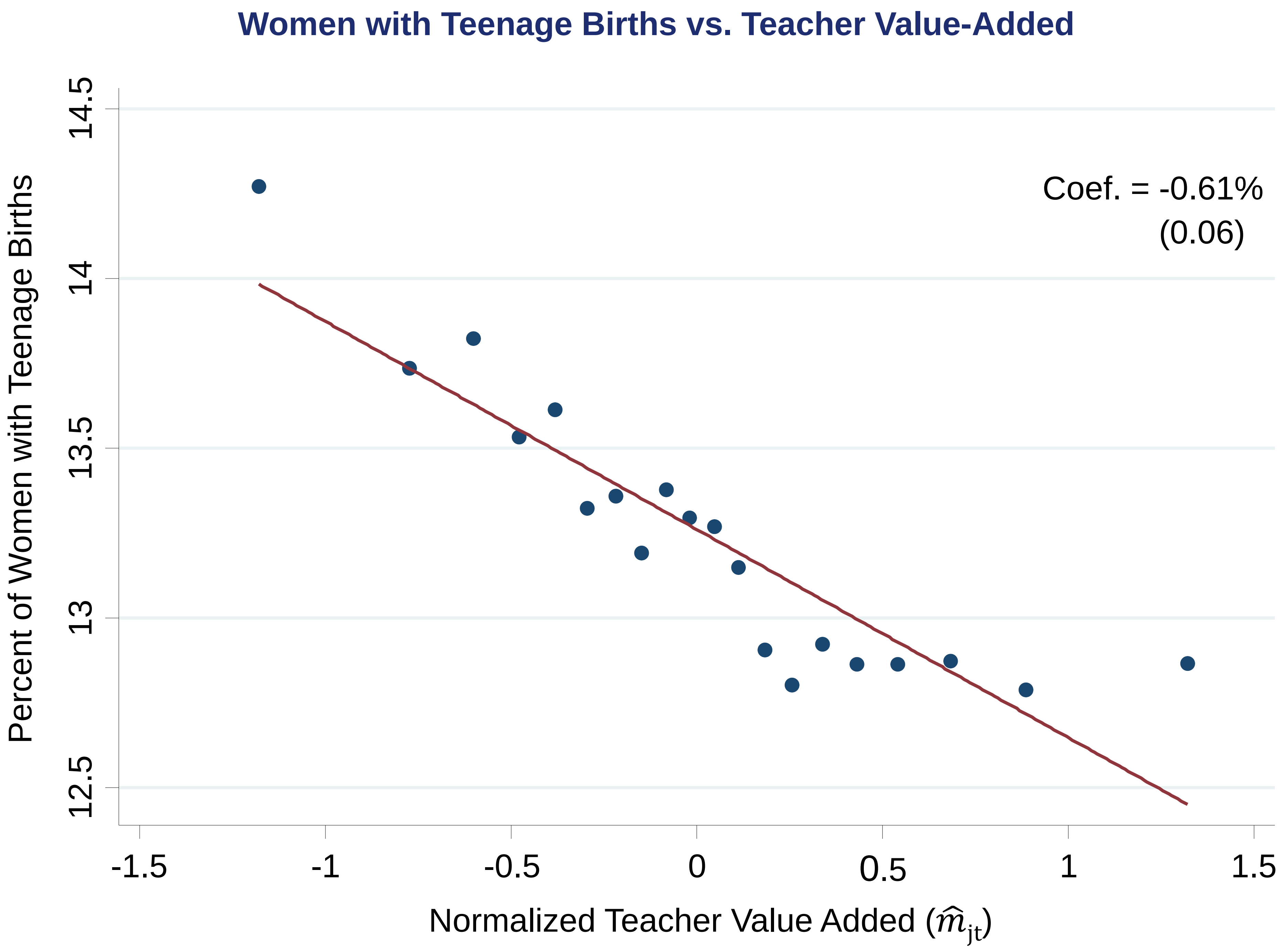

Earnings at Age 28 vs. Teacher Value-Added
Earnings at Age 28 ($)
22000
21500
21000
20500
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\hat{m}_{jt}$)
Coef. = $350
(92)

Impact of Teacher Value-Added on Earnings by Age

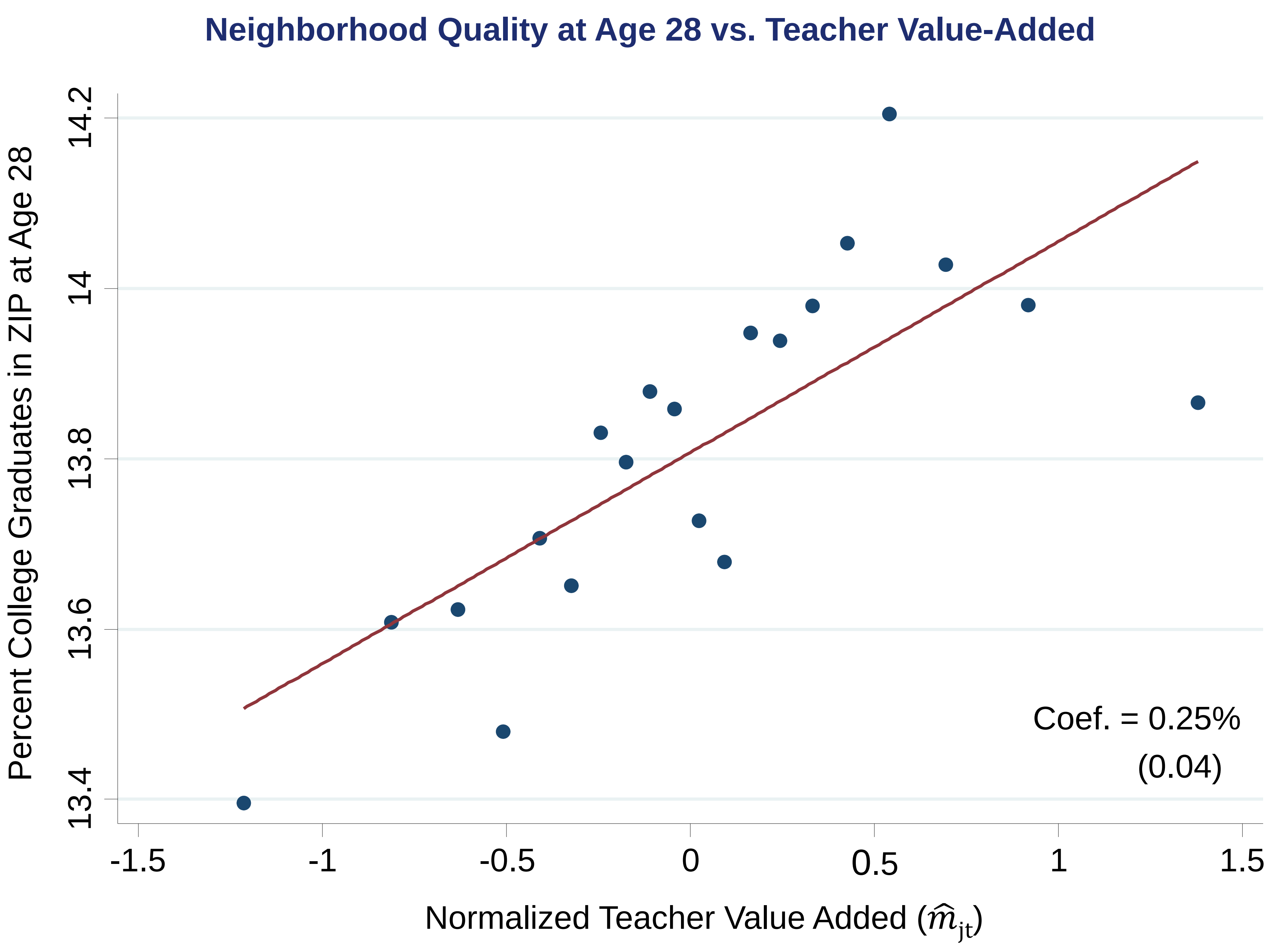


Impacts of Teacher Value-Added on Earnings

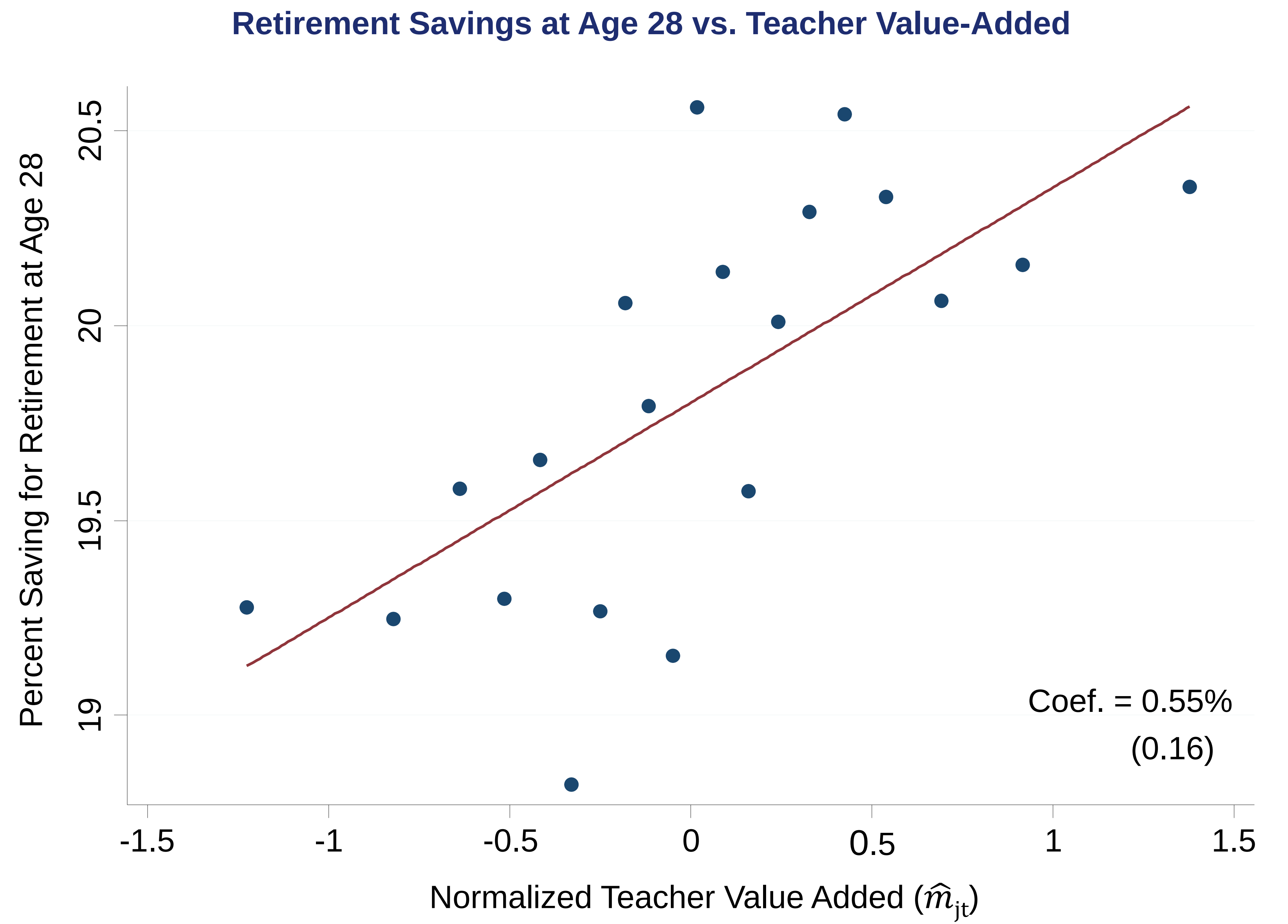
Dependent Variable:	Earnings at Age 28	Earnings at Age 28	Earnings at Age 28	Working at Age 28	Total Income at Age 28	Wage growth Ages 22-28
	($)	($)	($)	(%)	($)	($)
	(1)	(2)	(3)	(4)	(5)	(6)
Teacher VA	349.84	285.55	308.98	0.38	353.83	286.20
	(91.92)	(87.64)	(110.17)	(0.16)	(88.62)	(81.86)
Mean of Dep. Var.	21,256	21,256	21,468	68.09	22,108	11,454
Baseline Controls	X	X	X	X	X	X
Parent Chars. Controls		X			X	
Lagged Score Controls			X			
Observations	650,965	650,965	510,309	650,965	650,965	650,943



Women with Teenage Births vs. Teacher Value-Added
Percent of Women with Teenage Births
14.5
14
13.5
13
12.5
Coef. = -0.61%
(0.06)
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\hat{m}_{jt}$)



Neighborhood Quality at Age 28 vs. Teacher Value-Added
Percent College Graduates in ZIP at Age 28
14.2
14
13.8
13.6
13.4
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\widehat{m}_{jt}$)
Coef. = 0.25%
(0.04)



Retirement Savings at Age 28 vs. Teacher Value-Added
Percent Saving for Retirement at Age 28
19
19.5
20
20.5
-1.5
-1
-0.5
0
0.5
1
1.5
Normalized Teacher Value Added ($\hat{m}_{jt}$)
Coef. = 0.55%
(0.16)

Heterogeneity in Impacts of 1 SD of Teacher VA by Demographic Group

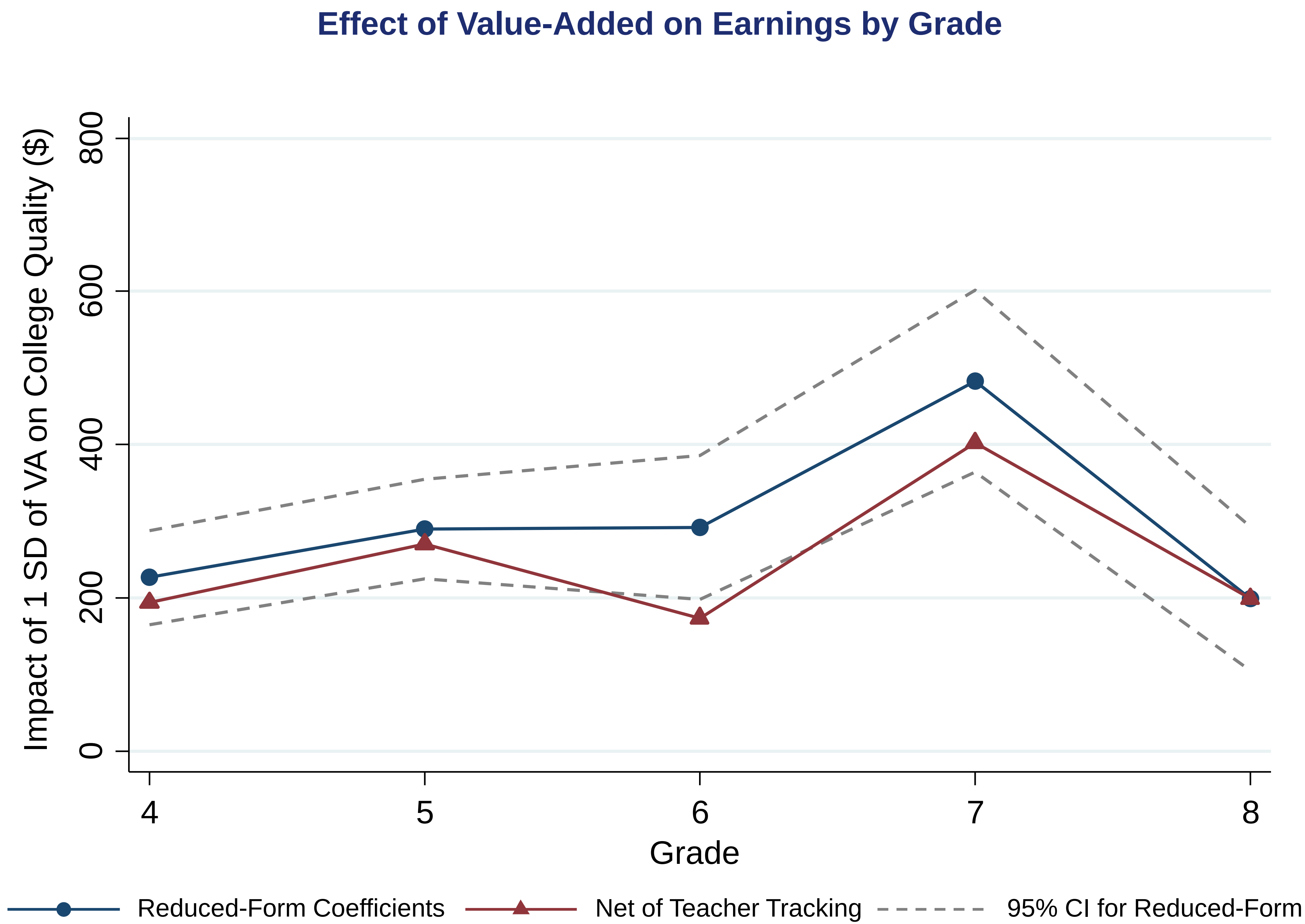
Dependent Variable:	College Quality at Age 20 ($)					
	Girls	Boys	Low Income	High Income	Minority	Non-Minority
	(1)	(2)	(3)	(4)	(5)	(6)
Value-Added	290.65 (23.61)	237.93 (21.94)	190.24 (19.63)	379.89 (27.03)	215.51 (17.09)	441.08 (42.26)
Mean College Quality	27,584	26,073	23,790	30,330	23,831	33,968
Impact as % of Mean	1.05%	0.91%	0.80%	1.25%	0.90%	1.30%

Heterogeneity in Impacts of 1 SD of Teacher VA by Subject

Dependent Variable:	College Quality at Age 20 ($)				
	Elementary School			Middle School	
	(1)	(2)	(3)	(4)	(5)
Math Teacher Value-Added	207.81 (21.77)		106.34 (28.50)	265.59 (43.03)	
English Teacher Value-Added		258.16 (25.42)	189.24 (33.07)		521.61 (63.67)
Control for Average VA in Other Subject				X	X

Teacher Impacts by Grade

- Reduced-form impacts of having better teachers in each grade include tracking to better teachers in future grades
- We can net-out the impact of tracking from the reduced-form coefficients by estimating tracking process
  - Estimate impact of current teacher VA on VA of future teachers
  - Subtract out impacts of future teachers



Effect of Value-Added on Earnings by Grade
Impact of 1 SD of VA on College Quality ($)
0
200
400
600
800
4
5
6
7
8
Grade
Reduced-Form Coefficients
Net of Teacher Tracking
95% CI for Reduced-Form

Policy Implications

Policy Proposal 1: Deselection of Low VA Teachers

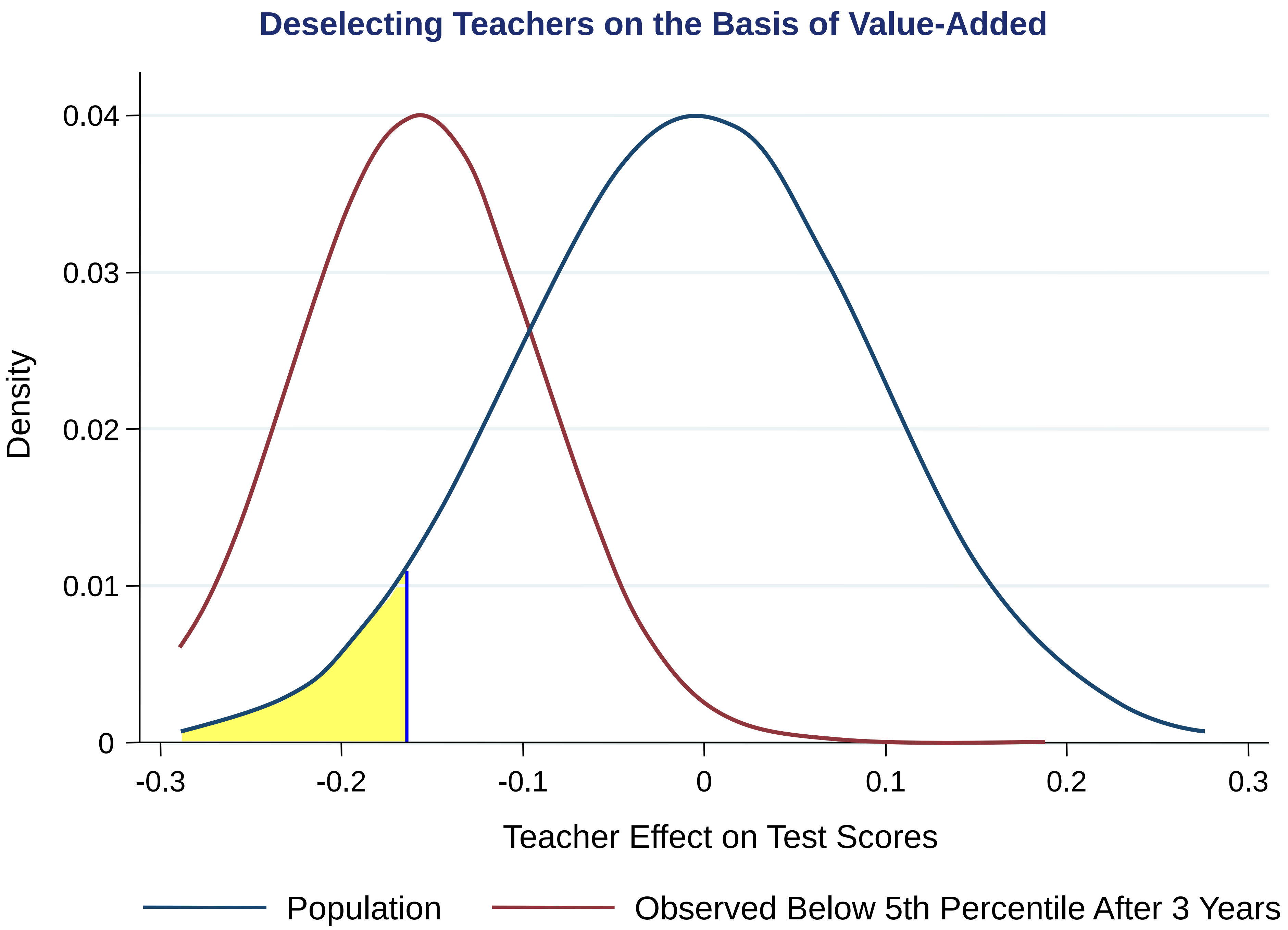
What are the gains from replacing teachers with VA in bottom 5% with teachers of median quality (Hanushek 2009)?

Policy Calculations

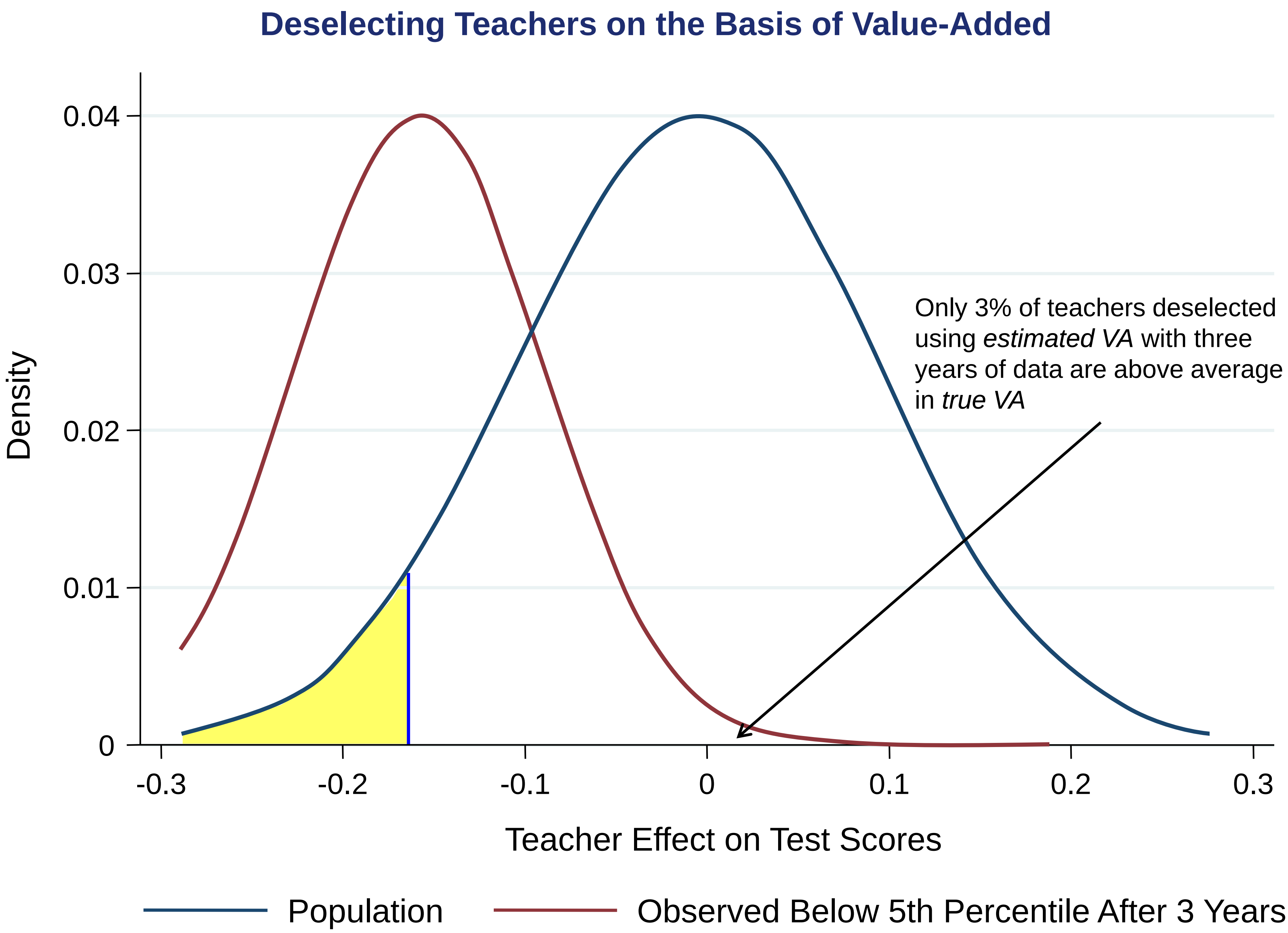
- Use estimates to evaluate gains from improving teacher quality
- Measure impact of teacher VA on present value of lifetime earnings
- Assumptions
  - Ignore general equilibrium effects and non-monetary gains
    [Oreopoulos and Salvanes 2011, Heckman 2000]
  - Constant percentage impact on earnings over life
  - Life-cycle earnings follows cross-sectional life-cycle path in 2010
  - 2% wage growth with 5% discount rate back to age 12
    - Undiscounted lifetime earnings gains are roughly 5 times larger

Policy Calculations

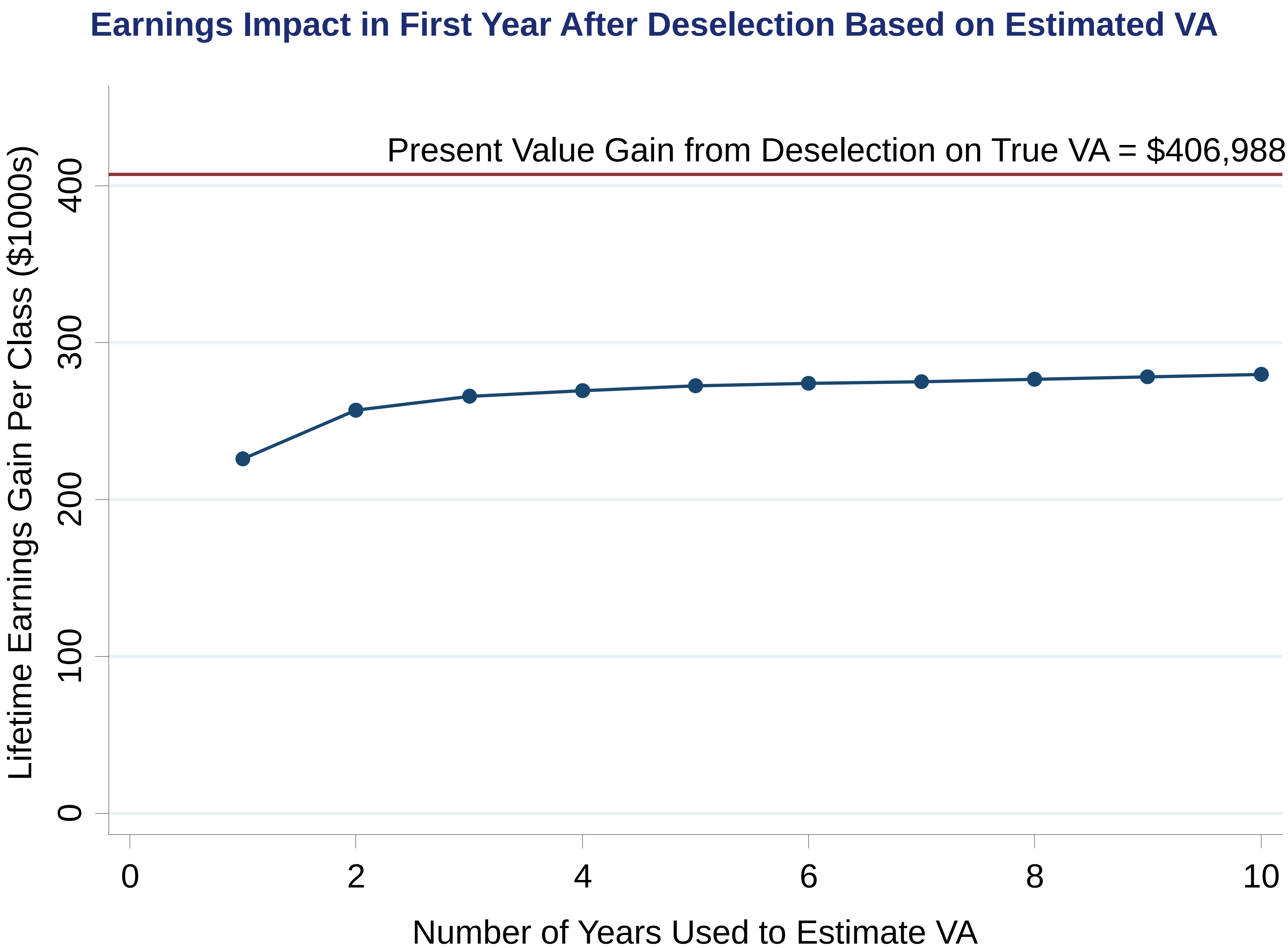
- Consider replacing teachers in the bottom 5% of VA distribution with teachers of average quality (Hanushek 2009)
- Select on *true* VA → NPV gain for a class of average size: $407,000
- In practice, gains are reduced by two factors
  - Estimation error in VA
  - Drift in VA over time



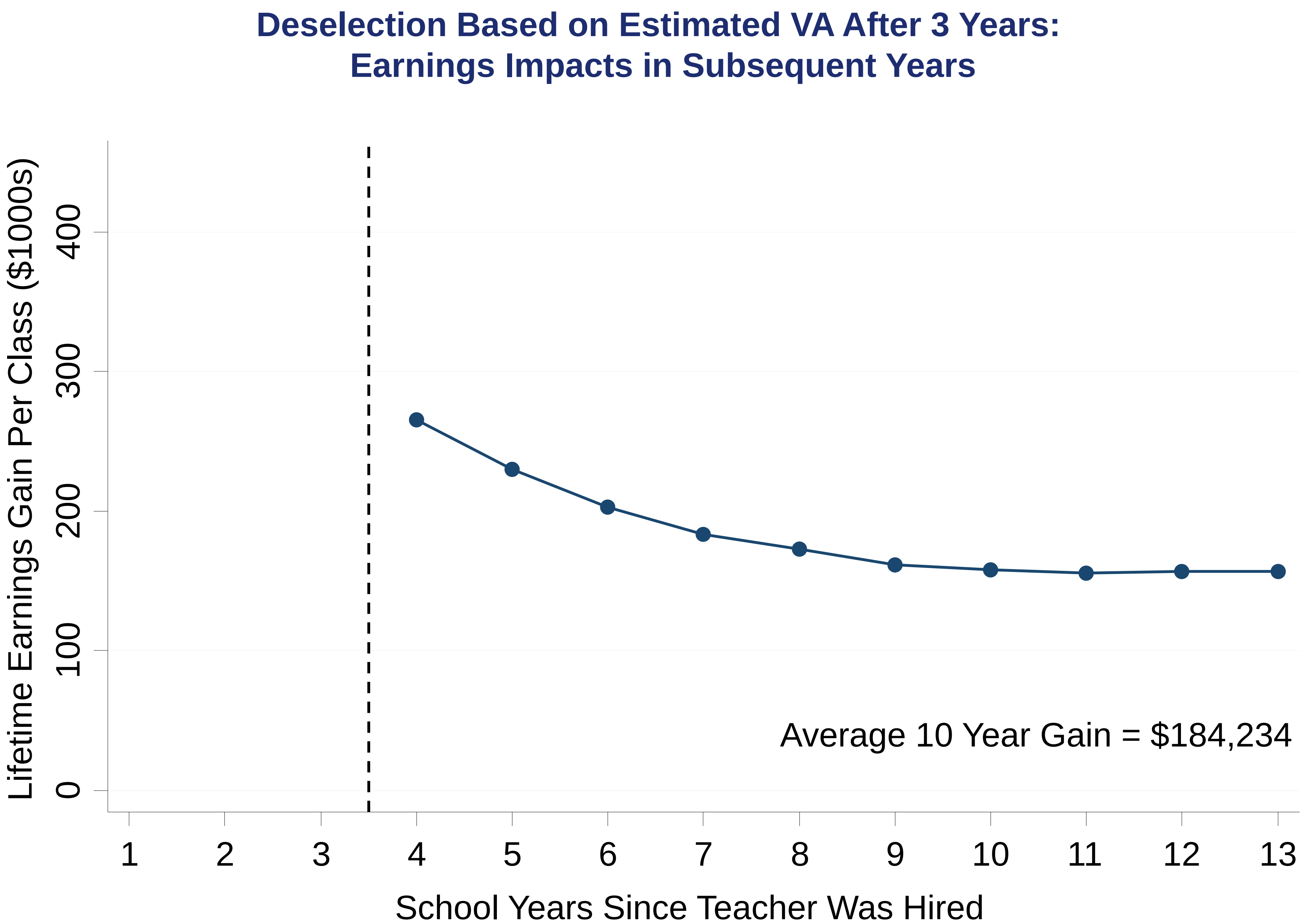
Deselecting Teachers on the Basis of Value-Added
Density
0.04
0.03
0.02
0.01
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Effect on Test Scores
Population
Observed Below 5th Percentile After 3 Years



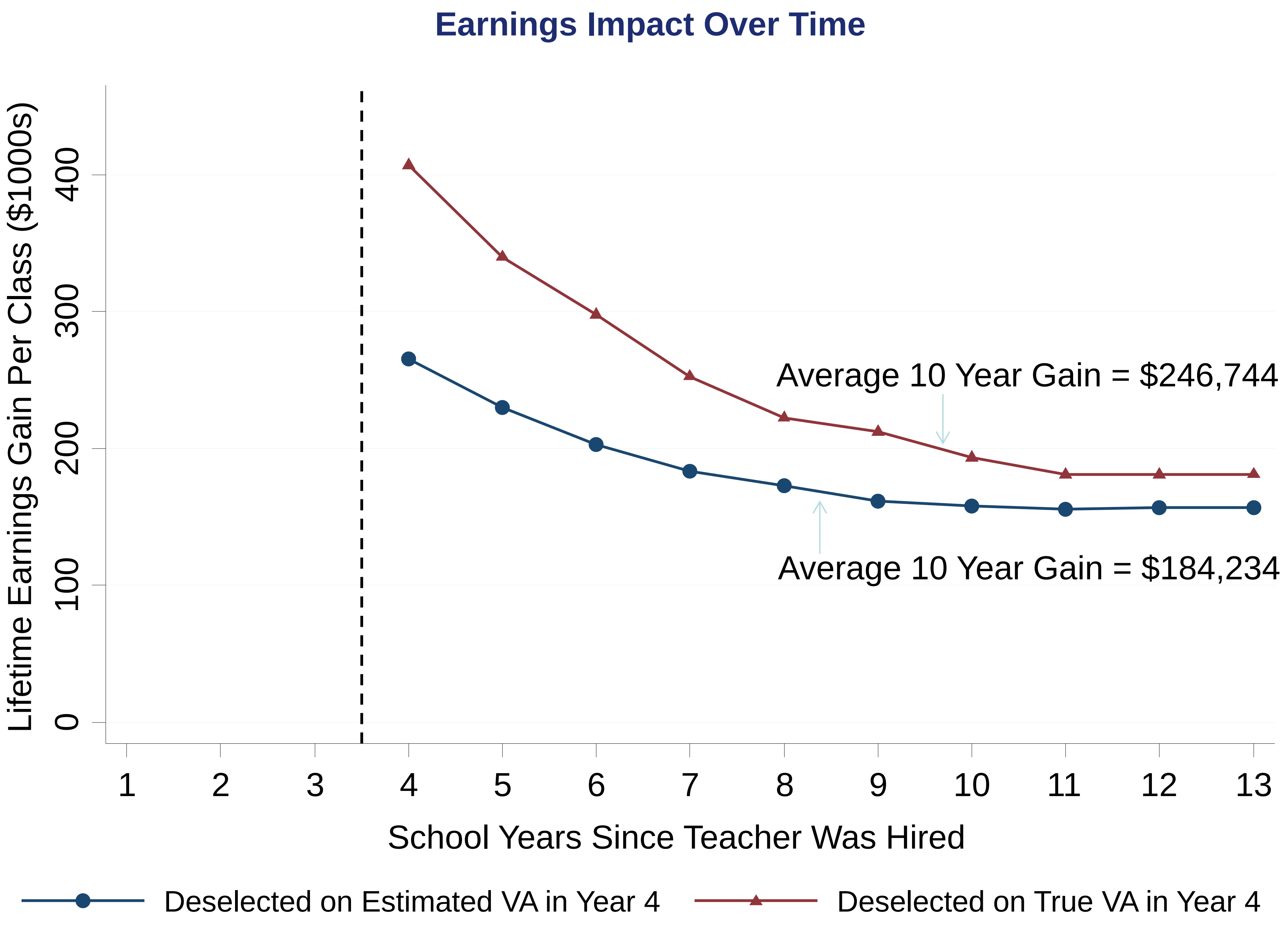
Deselecting Teachers on the Basis of Value-Added
Density
0.04
0.03
0.02
0.01
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Teacher Effect on Test Scores
Only 3% of teachers deselected using *estimated VA* with three years of data are above average in *true VA*
Population
Observed Below 5th Percentile After 3 Years



Earnings Impact in First Year After Deselection Based on Estimated VA
Present Value Gain from Deselection on True VA = $406,988
Lifetime Earnings Gain Per Class ($1000s)
0
100
200
300
400
0
2
4
6
8
10
Number of Years Used to Estimate VA



Deselection Based on Estimated VA After 3 Years:
Earnings Impacts in Subsequent Years
Lifetime Earnings Gain Per Class ($1000s)
0
100
200
300
400
1
2
3
4
5
6
7
8
9
10
11
12
13
School Years Since Teacher Was Hired
Average 10 Year Gain = $184,234



Earnings Impact Over Time
Lifetime Earnings Gain Per Class ($1000s)
0
100
200
300
400
1
2
3
4
5
6
7
8
9
10
11
12
13
School Years Since Teacher Was Hired
Average 10 Year Gain = $246,744
Average 10 Year Gain = $184,234
Deselected on Estimated VA in Year 4
Deselected on True VA in Year 4

Costs vs. Benefits of VA-Based Evaluation

- Rothstein (2013) estimates that deselecting bottom 5% of teachers based on VA would require a salary increase of $700 for all teachers

- Avg. gain from deselection policy is $184,000 x 5% = $9,250

- Gain 10 times as large as cost → VA could be a useful policy tool

- Key concern: gains may be eroded when VA is actually used
  - Using VA in high-stakes evaluation could lead to teaching to the test or cheating [Jacob 2005, Neal and Schanzenbach 2010, Barlevy and Neal 2012]

- Broader policy lesson: improving teacher quality, whether through VA or other metrics, likely to have very large returns

Policy Implications

Policy Proposal 2: Retention of High VA Teachers

What are the gains from increasing retention of high value-added teachers by paying salary bonuses?

Gains from Retaining High VA Teachers

- Retaining a teacher whose VA is at the 95th percentile (based on 3 years of data) for an extra year would generate PV earnings gain of **$266K**

- Clotfelter et al. (2008) analyze impacts of bonus payments to teachers

  - $1,800 bonus would raise teacher retention by 1.5 percentage points → earnings gain of **$3,200**

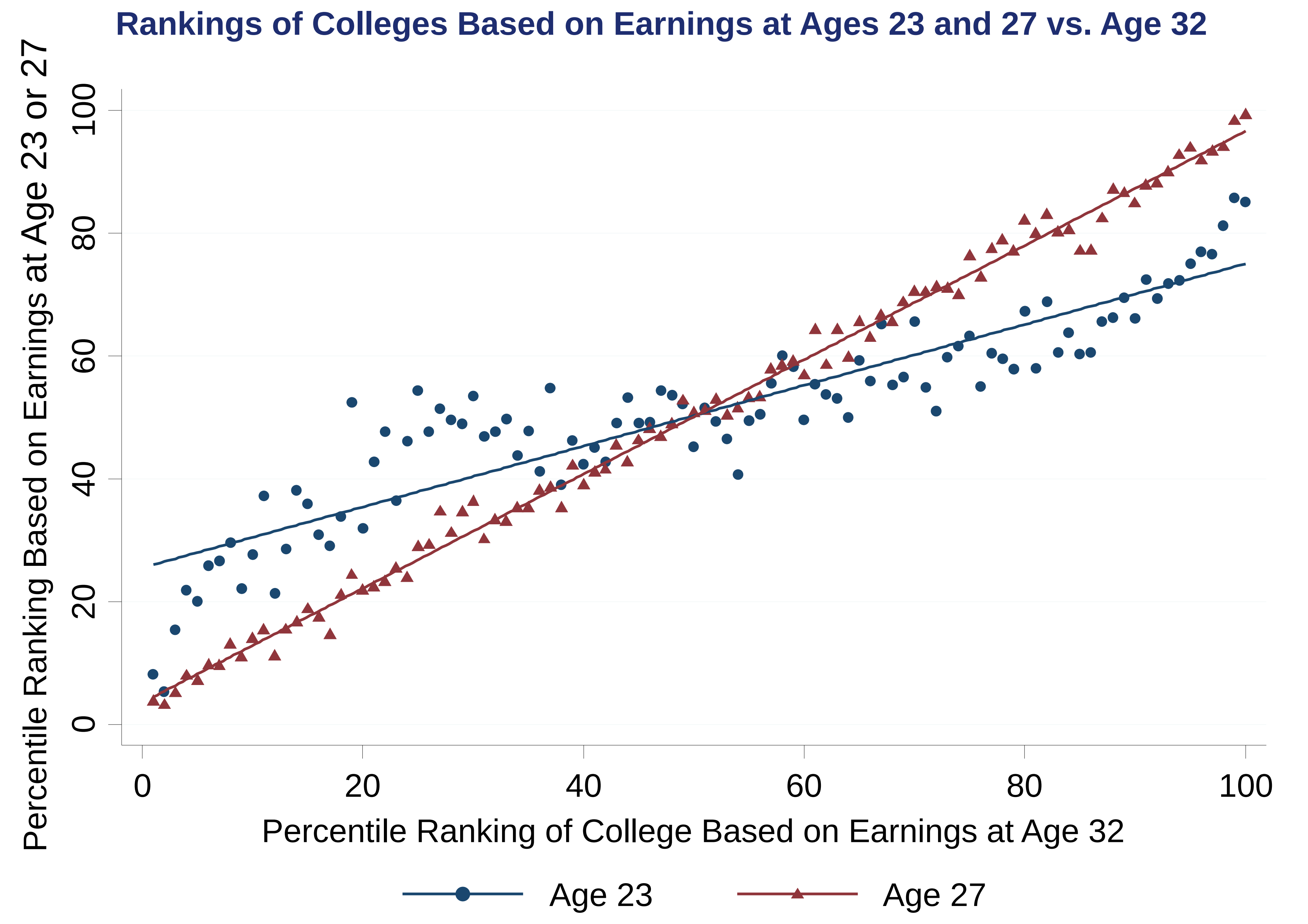
- Net return relatively small because most of the bonus payments go to teachers who would not have left anyway

  - Have to pay bonuses to 60 teachers to retain 1 teacher on average

Conclusion

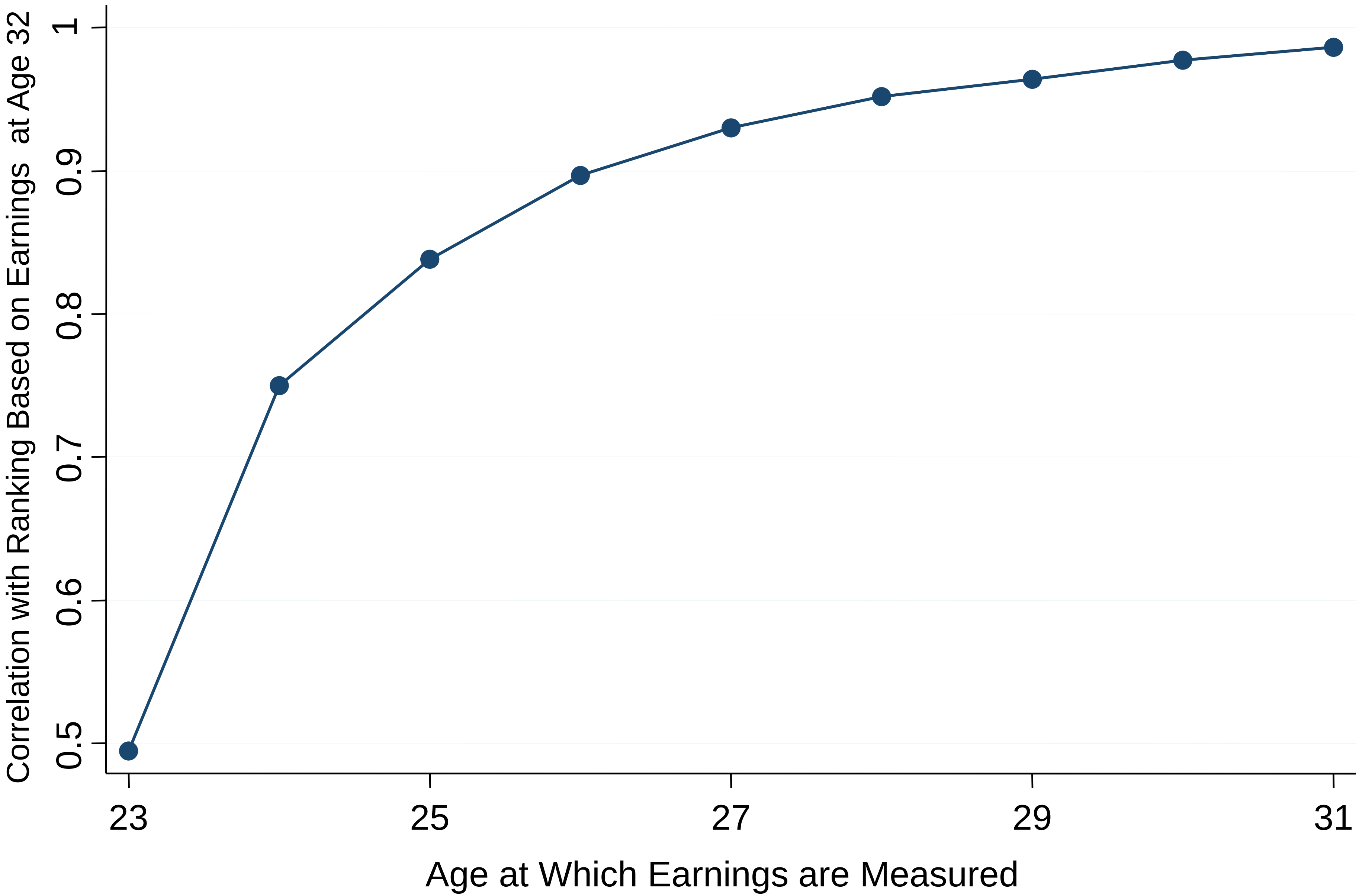
- Further work needed to assess value-added as a policy tool
  - Using VA measures in high-stakes evaluation could induce negative behavioral responses such as teaching to the test or cheating
  - Errors in personnel decisions must be weighed against mean benefits
- Results highlight large potential returns from developing policies to improve teacher quality
- From a purely financial perspective, parents should be willing to pay about **$7,000/year** to get a 1 SD higher VA teacher for their child

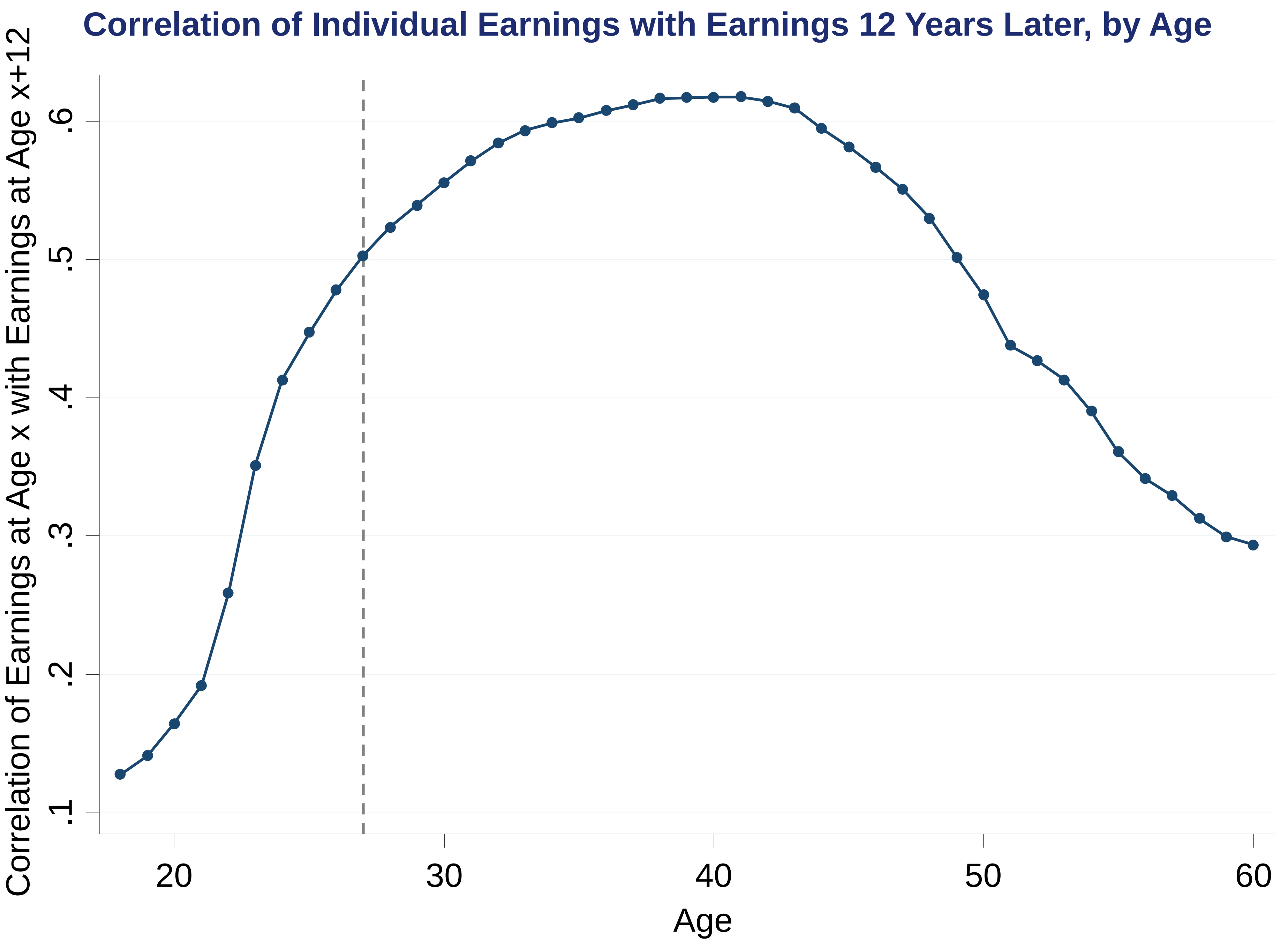
Appendix Figures

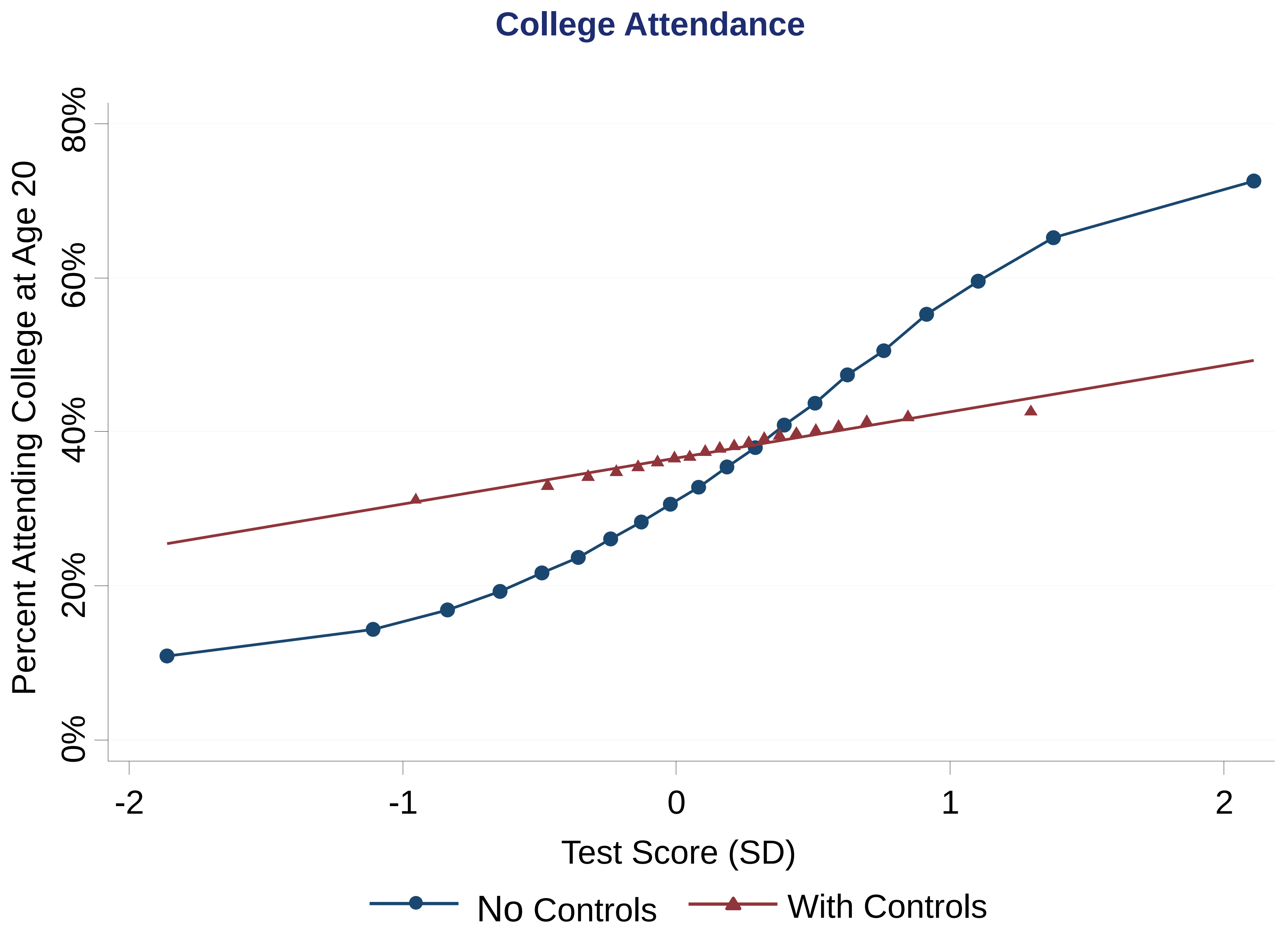


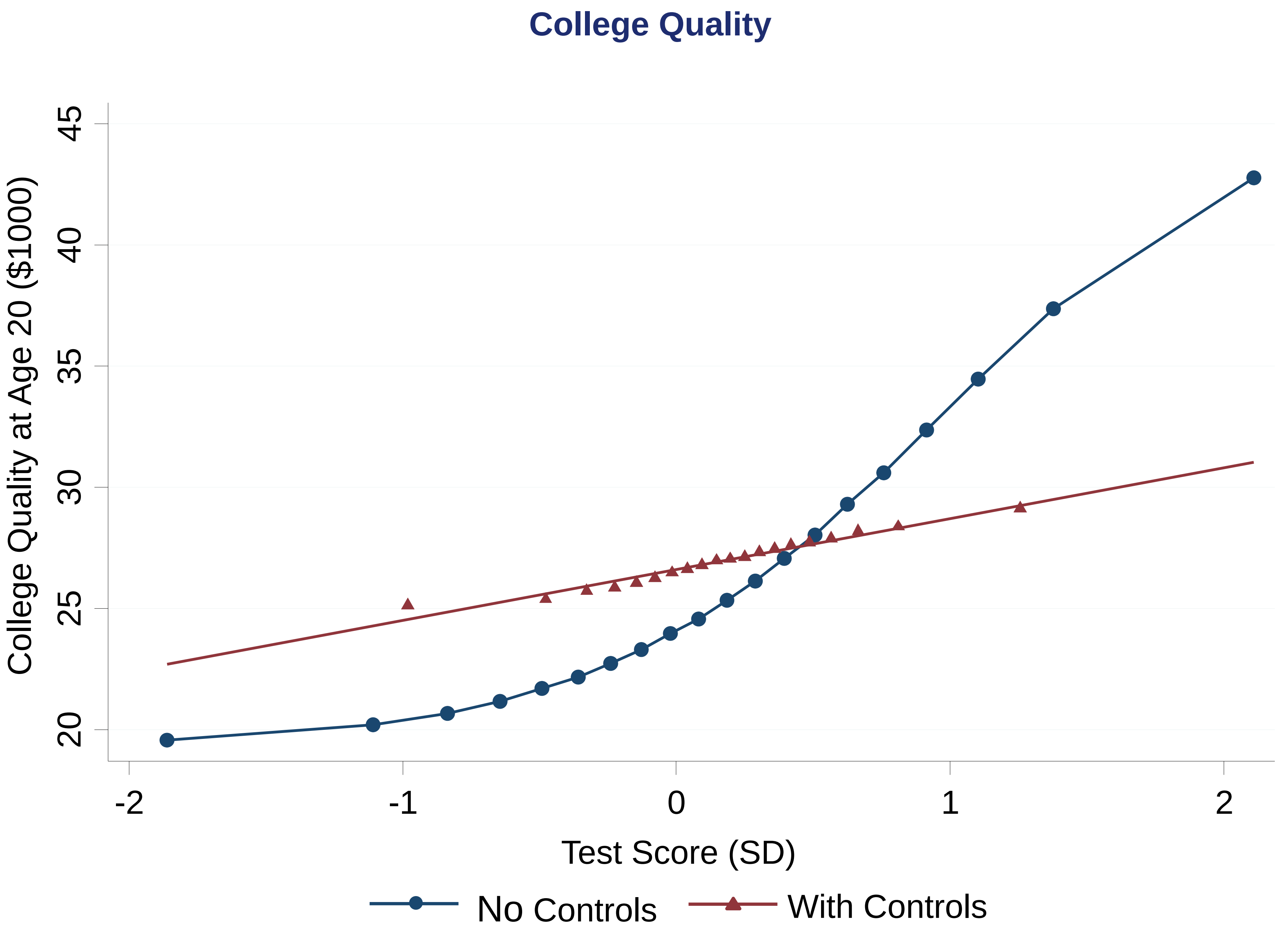
Rankings of Colleges Based on Earnings at Ages 23 and 27 vs. Age 32
Percentile Ranking Based on Earnings at Age 23 or 27
0
20
40
60
80
100
Percentile Ranking of College Based on Earnings at Age 32
Age 23
Age 27

Correlation of College Rankings Based on Earnings at Age 32
With Rankings Based on Earnings at Earlier Ages

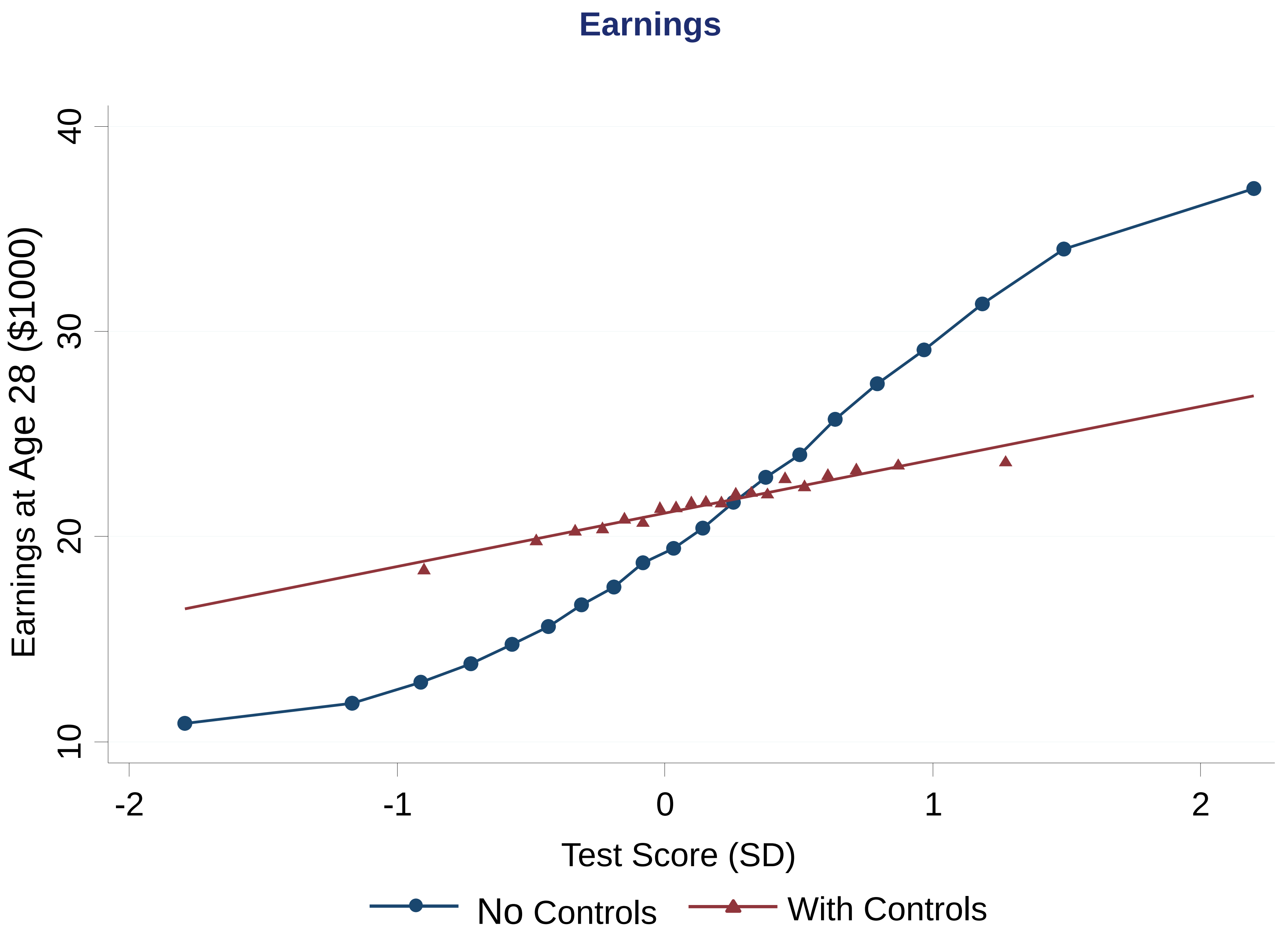



Correlation of Individual Earnings with Earnings 12 Years Later, by Age
Correlation of Earnings at Age x with Earnings at Age x+12
.1
.2
.3
.4
.5
.6
20
30
40
50
60
Age

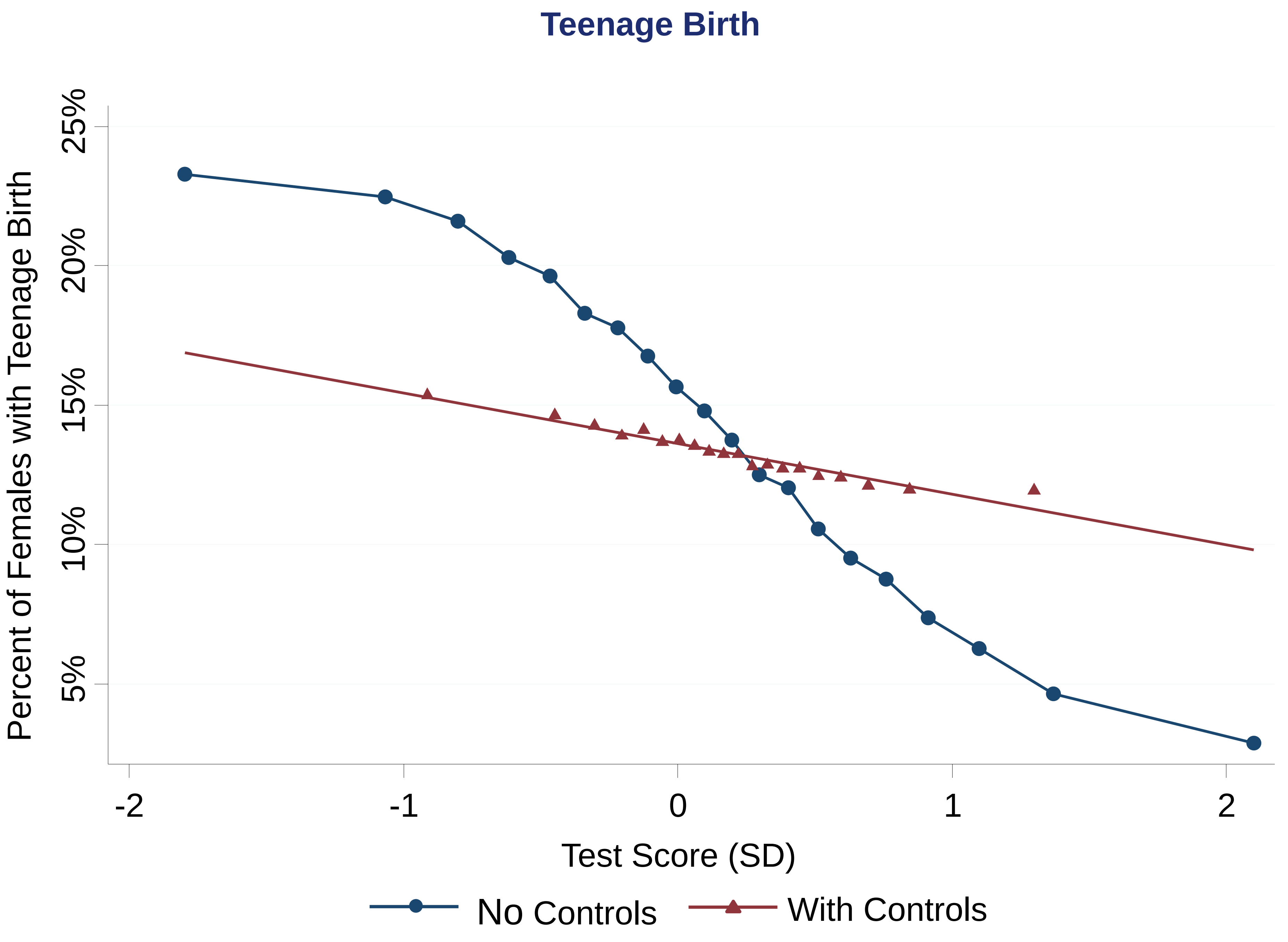

College Attendance
Percent Attending College at Age 20
0%
20%
40%
60%
80%
-2
-1
0
1
2
Test Score (SD)
No Controls
With Controls



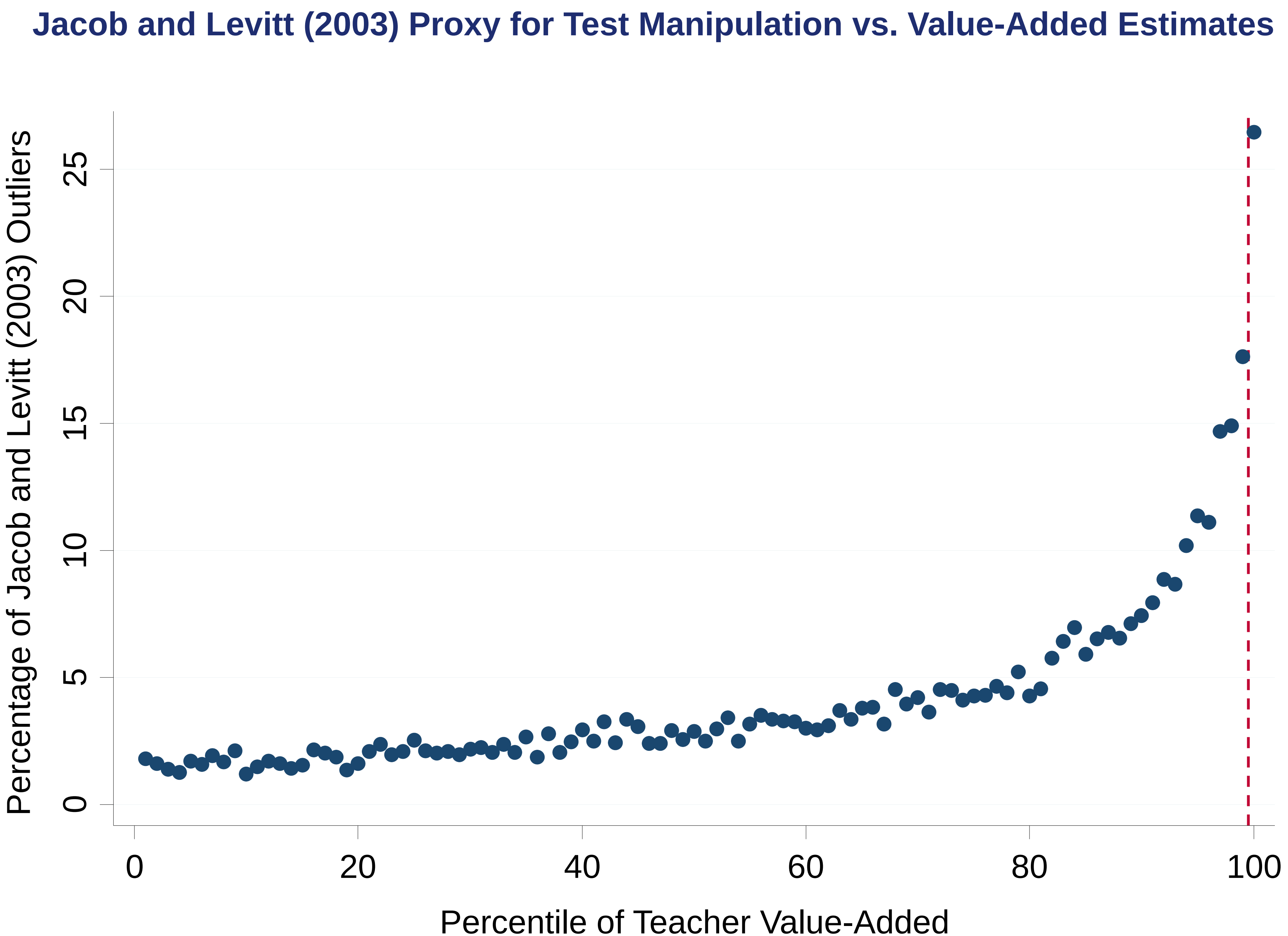
College Quality
College Quality at Age 20 ($1000)
20
25
30
35
40
45
-2
-1
0
1
2
Test Score (SD)
No Controls
With Controls



Earnings
Earnings at Age 28 ($1000)
10
20
30
40
-2
-1
0
1
2
Test Score (SD)
No Controls
With Controls



Teenage Birth
Percent of Females with Teenage Birth
25%
20%
15%
10%
5%
-2
-1
0
1
2
Test Score (SD)
No Controls
With Controls


Jacob and Levitt (2003) Proxy for Test Manipulation vs. Value-Added Estimates
Percentage of Jacob and Levitt (2003) Outliers
0
5
10
15
20
25
0
20
40
60
80
100
Percentile of Teacher Value-Added