Evaluating the Effect of New School Facilities on Student Achievement & Attendance in LAUSD

Julien Lafortune[1]
David Schönholzer[1]

[1]UC Berkeley, Department of Economics

BEAR Seminar, February 2017

Introduction: School Infrastructure Investments

- School infrastructure is an important component of K-12 spending:
  - $\Rightarrow$ \$45 billion spent on capital expenditures in US schools in 2012
  - $\Rightarrow$ \$13 billion spent in 2013 on school constructions
- Most research focused on effects of *instructional expenditures*, with less attention on capital expenditures
- School facilities are important component of public infrastructure, more generally
  - $\Rightarrow$ Potential bipartisan support for increasing infrastructure spending
  - $\Rightarrow$ Low interest rates – financing public works projects cheap

Motivation: New Facility Effects on Student Outcomes

1. Large disparities in school facility quality between rich and poor students, white and minority students, etc
2. No consensus in literature on impact of school capital expenditures on student outcomes
3. Little empirical work examining potential mechanisms

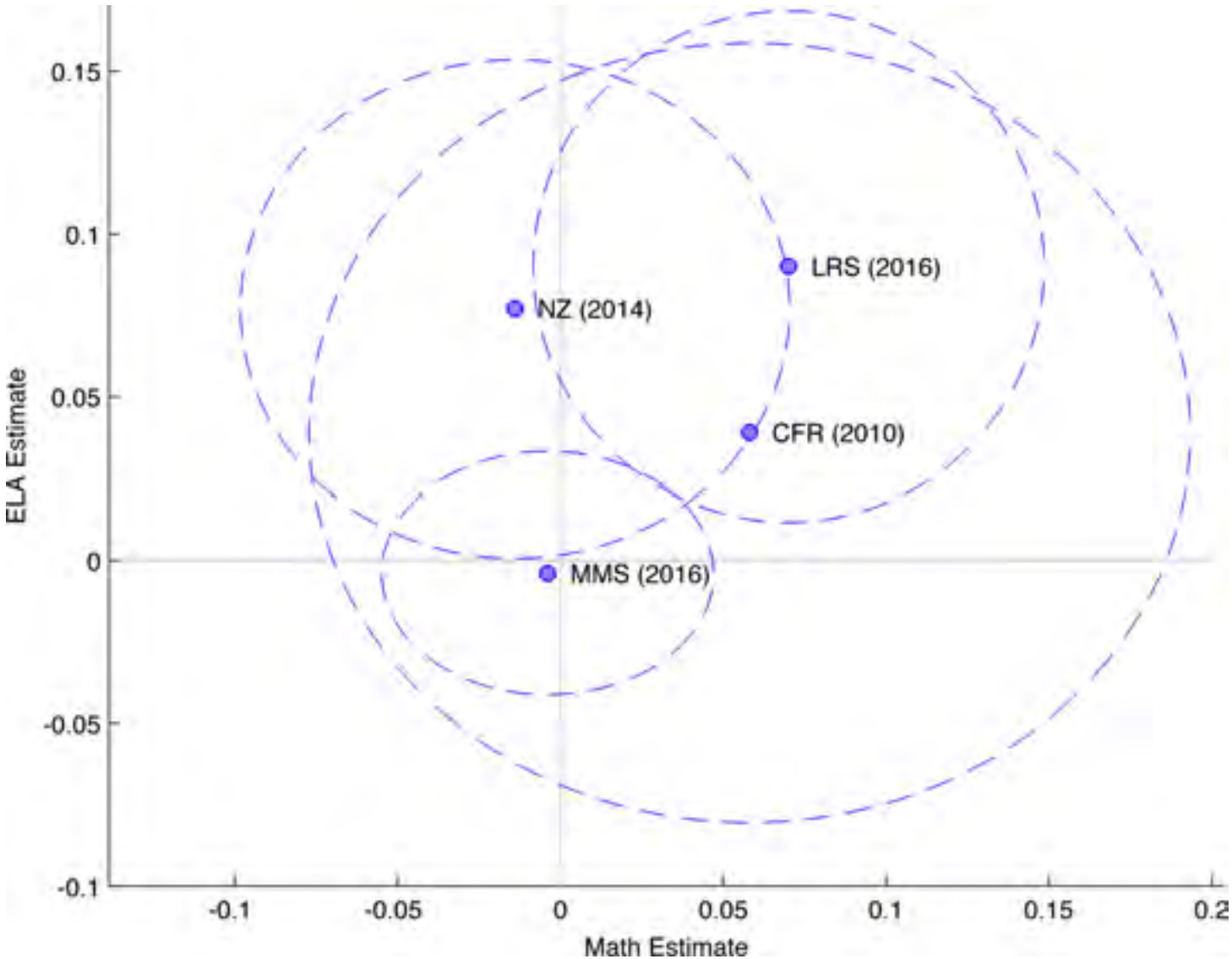
Research Question: What is the impact of new school constructions on student outcomes? What mechanisms might underly any effects?

This Paper

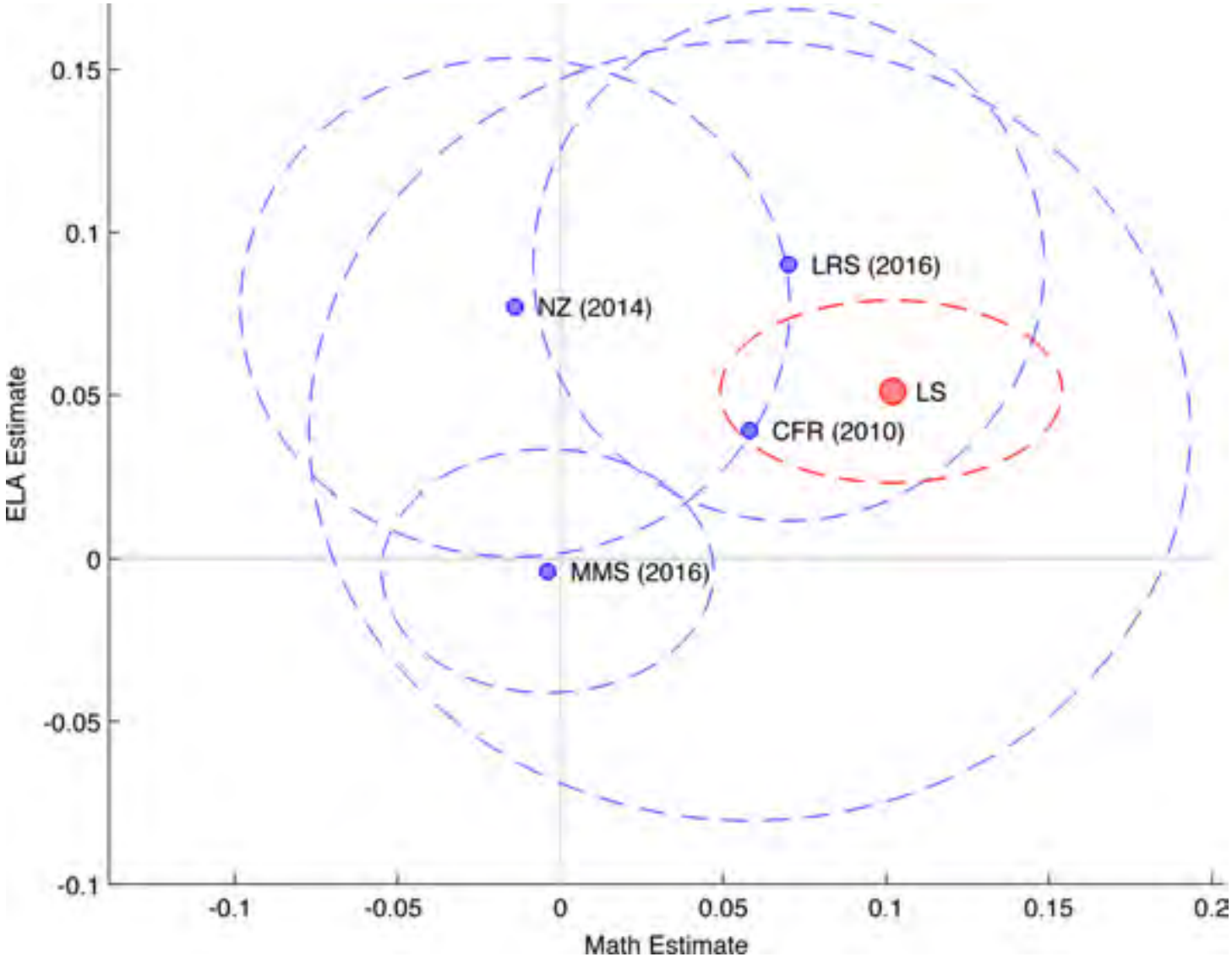
- Program evaluation of largest school construction program in US History:
  - $\Rightarrow$ Since 1998, Los Angeles Unfied School District (LAUSD) has allocated $27 billion dollars to capital expenditure programs (mainly state and local money)

- Exploit variation in timing and location of new school constructions to examine potential student-level impacts
  - $\Rightarrow$ Event study design around time student begins attending newly constructed school
  - $\Rightarrow$ Outcomes: student test scores (math, ELA) and attendance

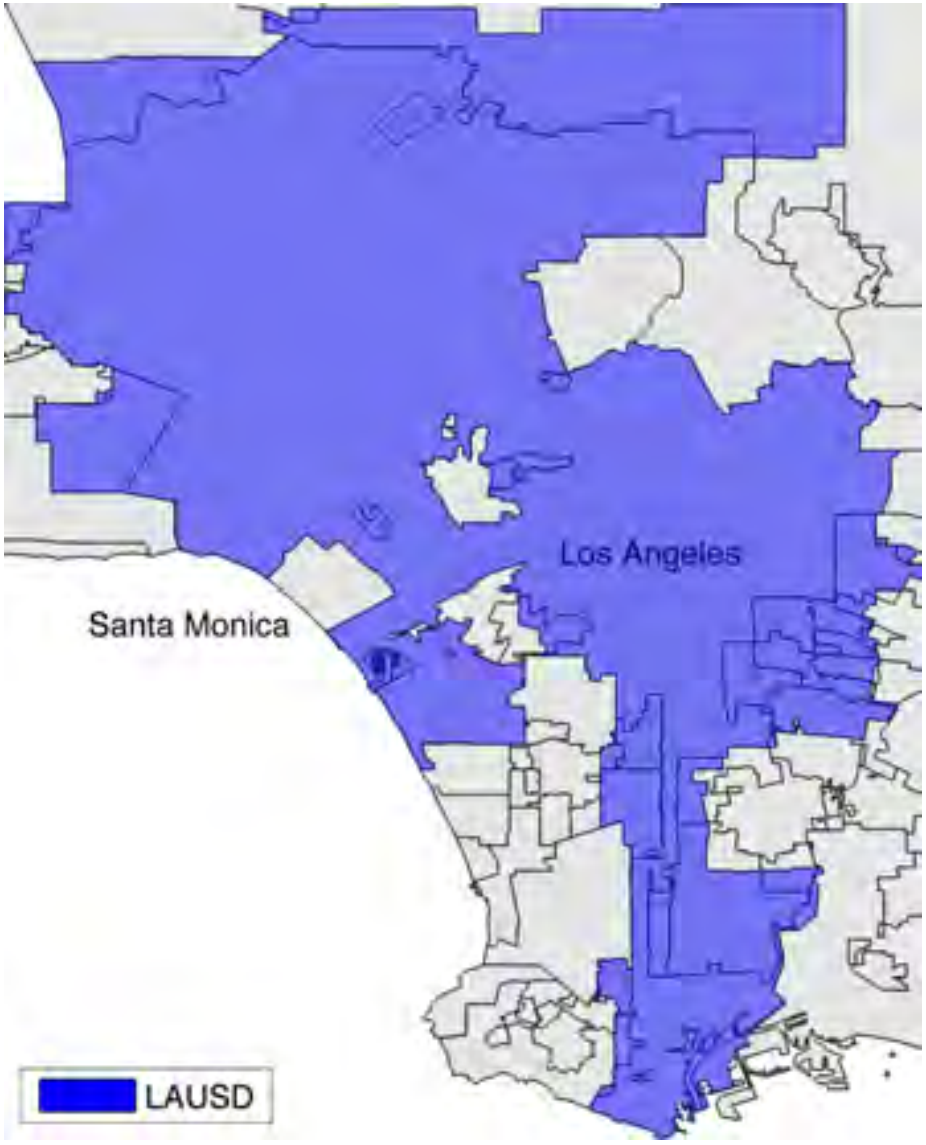
School Construction (Economics) Literature Estimates



Our Estimates



LAUSD in the L.A. Metro Area

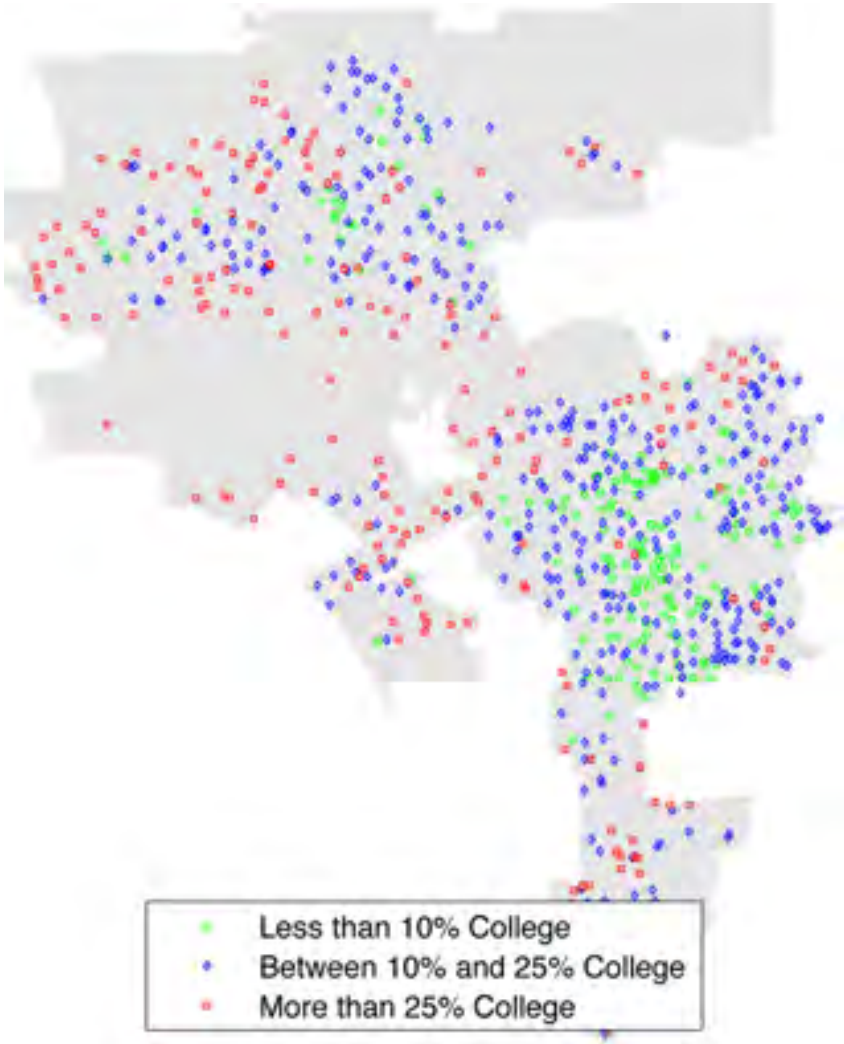


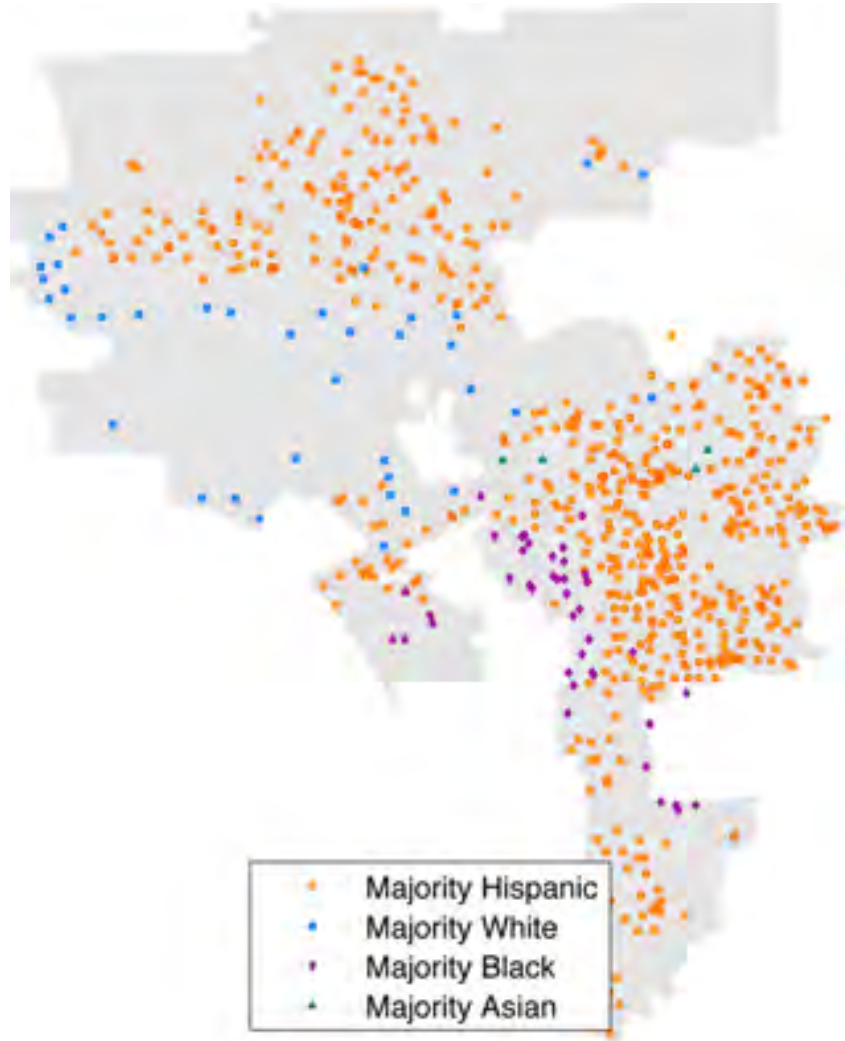
- 2nd largest district in U.S.
- 747,009 students at peak
- Mostly non-white district
- Serves 26 cities:
  - City of L.A.
  - Some gateway cities
  - Unincorporated areas
  - *Not* e.g. Santa Monica
- Underachieving:
  - -0.2 SD below CA in Math
  - -0.25 SD in ELA
- Lack of facility investment:

  $\Rightarrow$ Poor facility quality

  $\Rightarrow$ Overcrowding

LAUSD Socio-Demographics by School

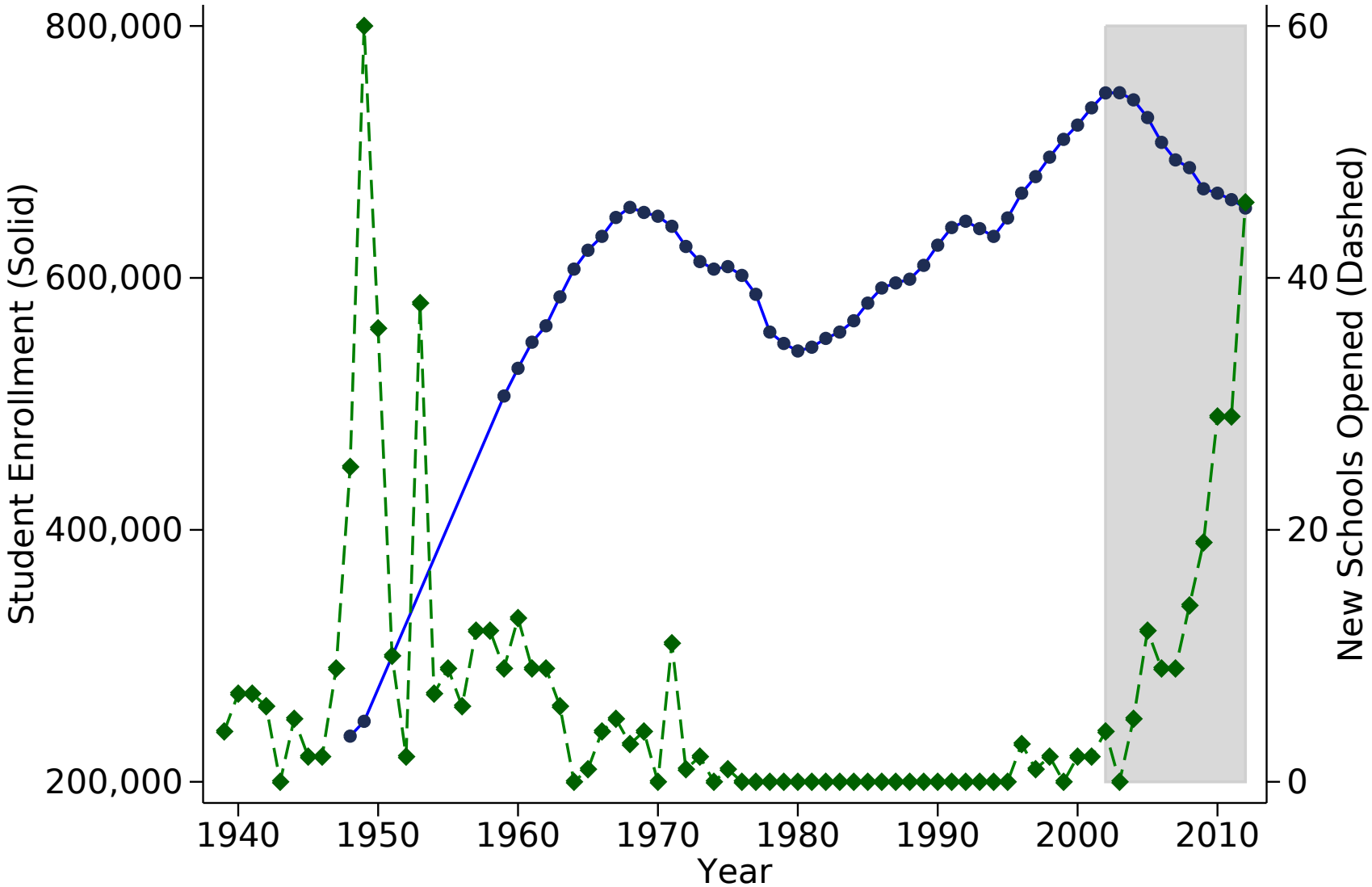




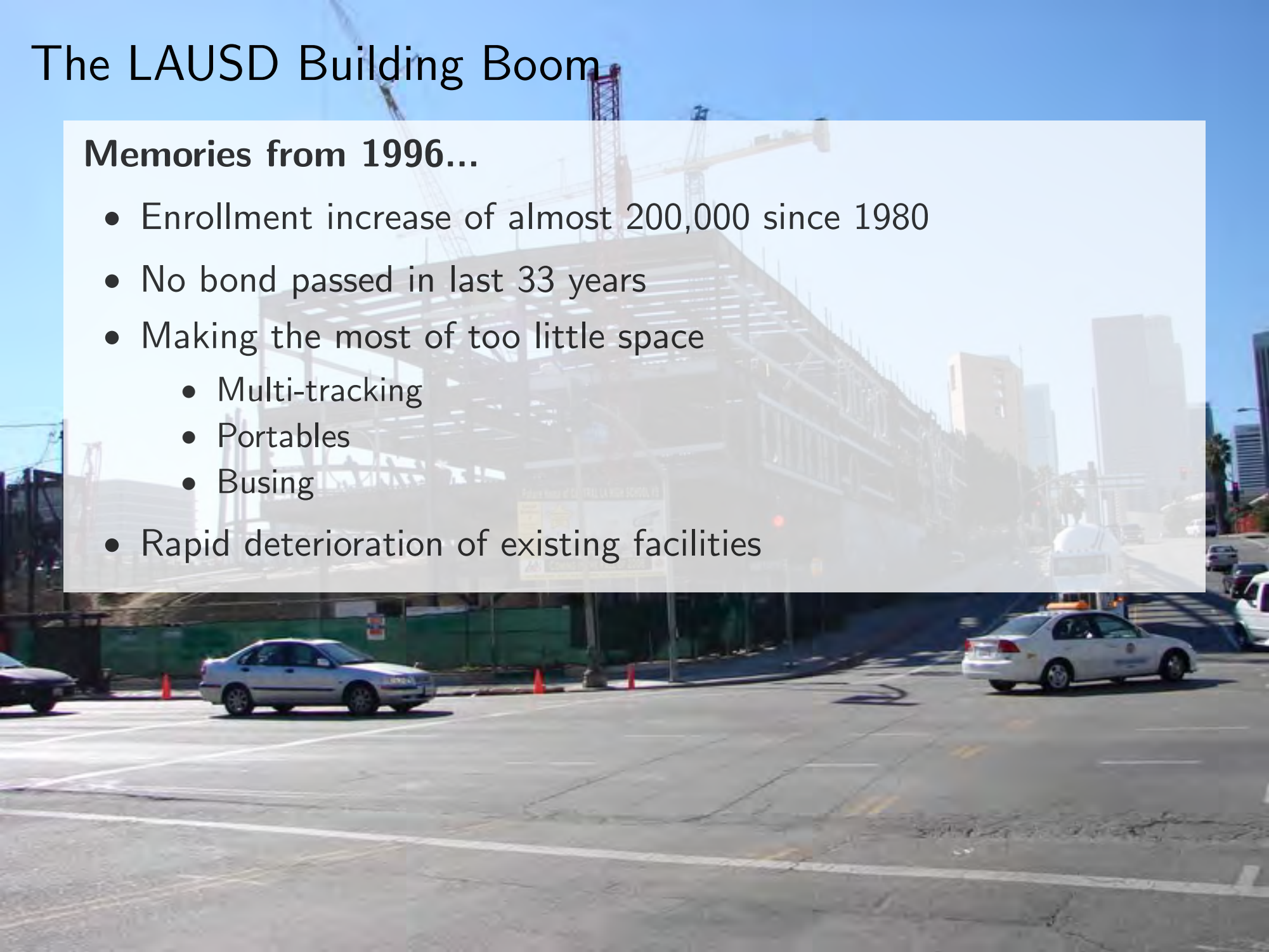
Section 1

Historical Context

School Construction and Enrollment 1940-2012



The LAUSD Building Boom



Memories from 1996...

- Enrollment increase of almost 200,000 since 1980
- No bond passed in last 33 years
- Making the most of too little space
  - Multi-tracking
  - Portables
  - Busing
- Rapid deterioration of existing facilities

The LAUSD Building Boom

Memories from 1996...

- Enrollment increase of almost 200,000 since 1980
- No bond passed since 33 years
- Making the most of too little space
  - Multi-tracking
  - Portables
  - Busing
- Rapid deterioration of existing facilities

Breakthrough

- 150 new schools built in 2002-2012
- About 150,000 new 2-semester seats
- Largest school building boom in U.S. history


92.5

Poor Quality Facilities

- Common facility quality issues:
  - Broken tables, blackboards, other teaching materials
  - Broken plumbing, ventilation, heating; closed bathrooms
  - Pest infestation, mold, mites
  - Lead paint and arsenic
- Anecdotal effects or poor facility conditions:
  - Temperature and noise distraction
  - Low student and teacher motivation
  - Health issues such as asthma and developmental disorders

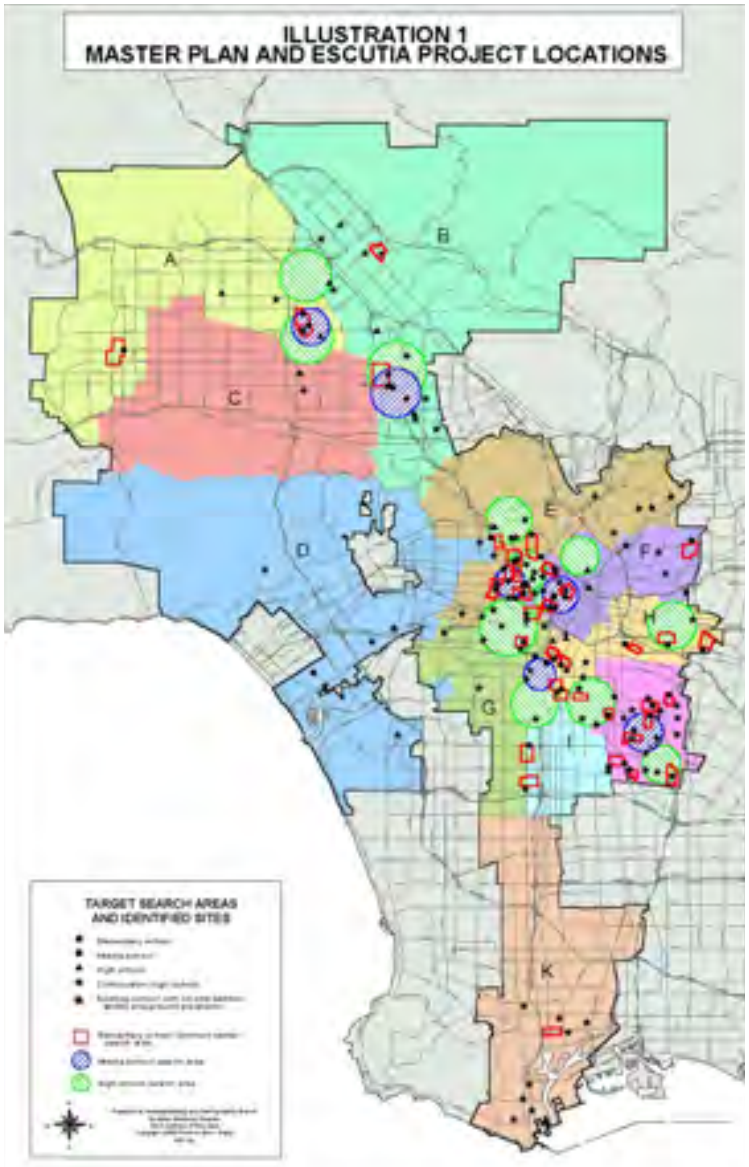
Overcrowding

- Common overcrowding conditions:
  - Temporary classrooms (portables)
  - Convert gyms, libraries, computer labs into classrooms
  - Multi-track calendars (year-round schools)
  - Long school ways, some busing (2-3%)
  - Overcrowded classrooms
- Anecdotal effects of overcrowding:
  - Diminished attention of students
  - Increased school violence
  - Limited access to non-classroom opportunities
  - Multi-track: longer school days and shorter school year
  - Rapid deterioration of facility conditions

New and Old School Sites in LAUSD

- School facility bonds:
  - 1997: \$2.4 billion
  - 2002: \$3.35 billion
  - 2004: \$3.87 billion
  - 2005: \$4 billion
  - 2007: \$7 billion
- Building boom 2002-2012:
  - $\Rightarrow$ 148 new schools
  - $\Rightarrow$ 19% increase
  - $\Rightarrow$ Higher facility standards

New School Site Selection Process



- Select old schools most...
  1. overcrowded
  2. multi-track calendar

  $\Rightarrow$ 109 schools identified (black dots)
- Assign search areas nearby:
  - Red: elementary schools
  - Blue: middle schools
  - Green: high schools
- Select sites from areas:
  - Feasibility study
  - CEQA
  - Property purchase
  - Public tender
  - Construction (1-3 years)

Relieved Overcrowding

Example: Madison Elementary

Example: Robert F. Kennedy Family of Schools

Computer Labs, Libraries, etc. Back to Original Purpose

Section 2

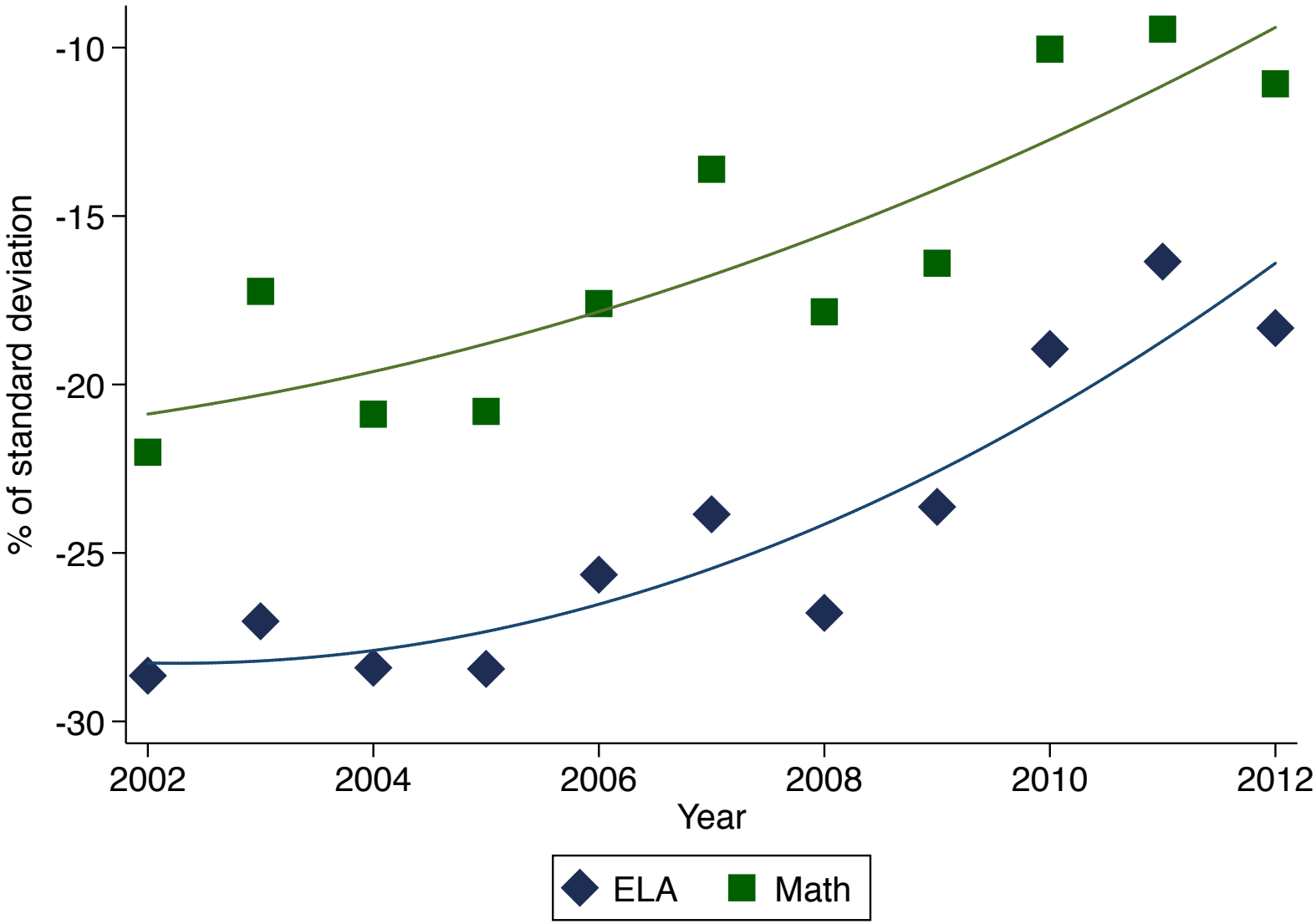
Data

Data

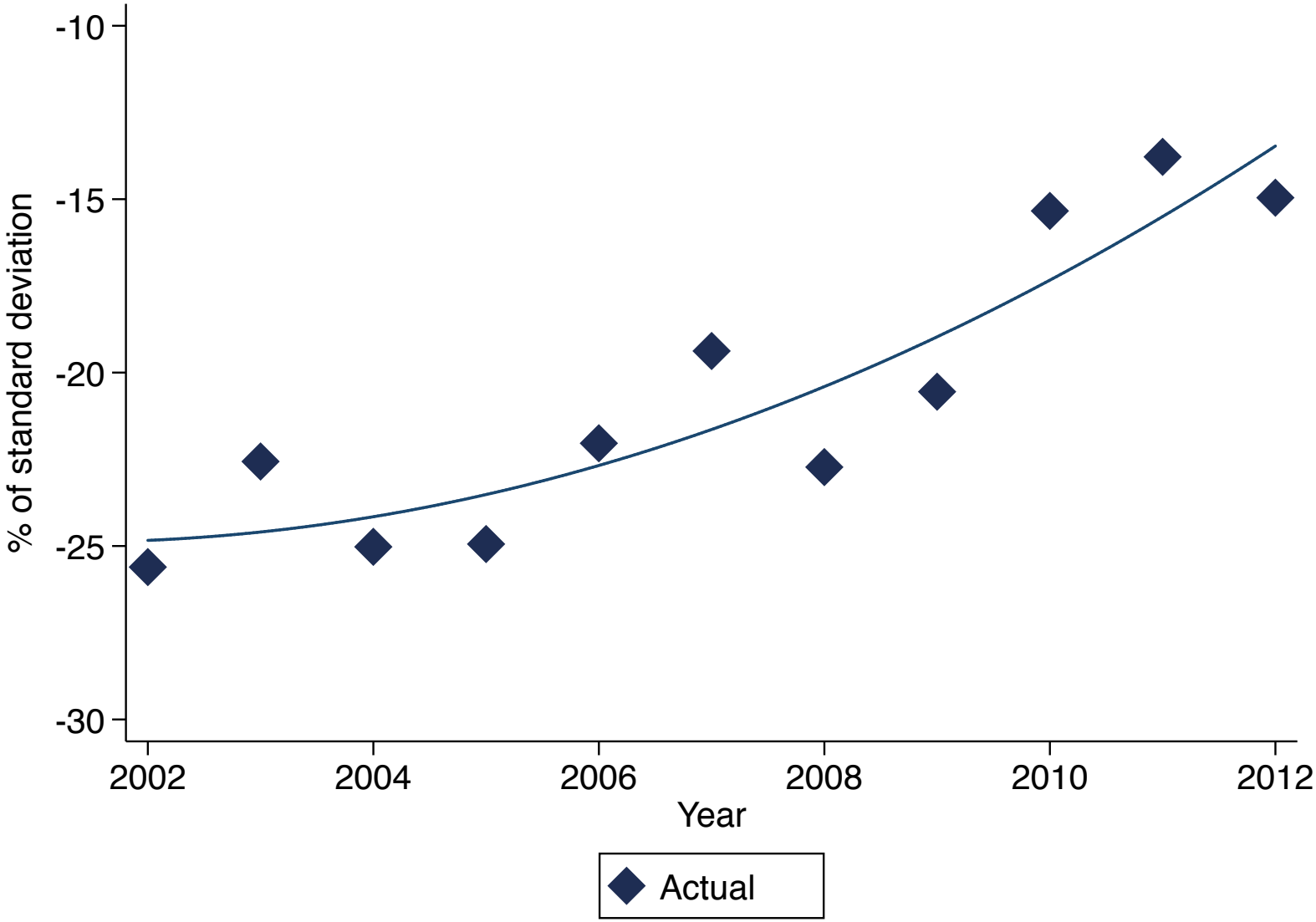
Two primary data sources:

1. Administrative data from LAUSD for 2002-2012
   - Math and ELA test scores G2-G11
   - Demographics
   - Attendance (annual)
   - Teacher records
2. New school projects from LAUSD Facilities Service Division
   - Location
   - Cost, number of seats
   - Completion timeline

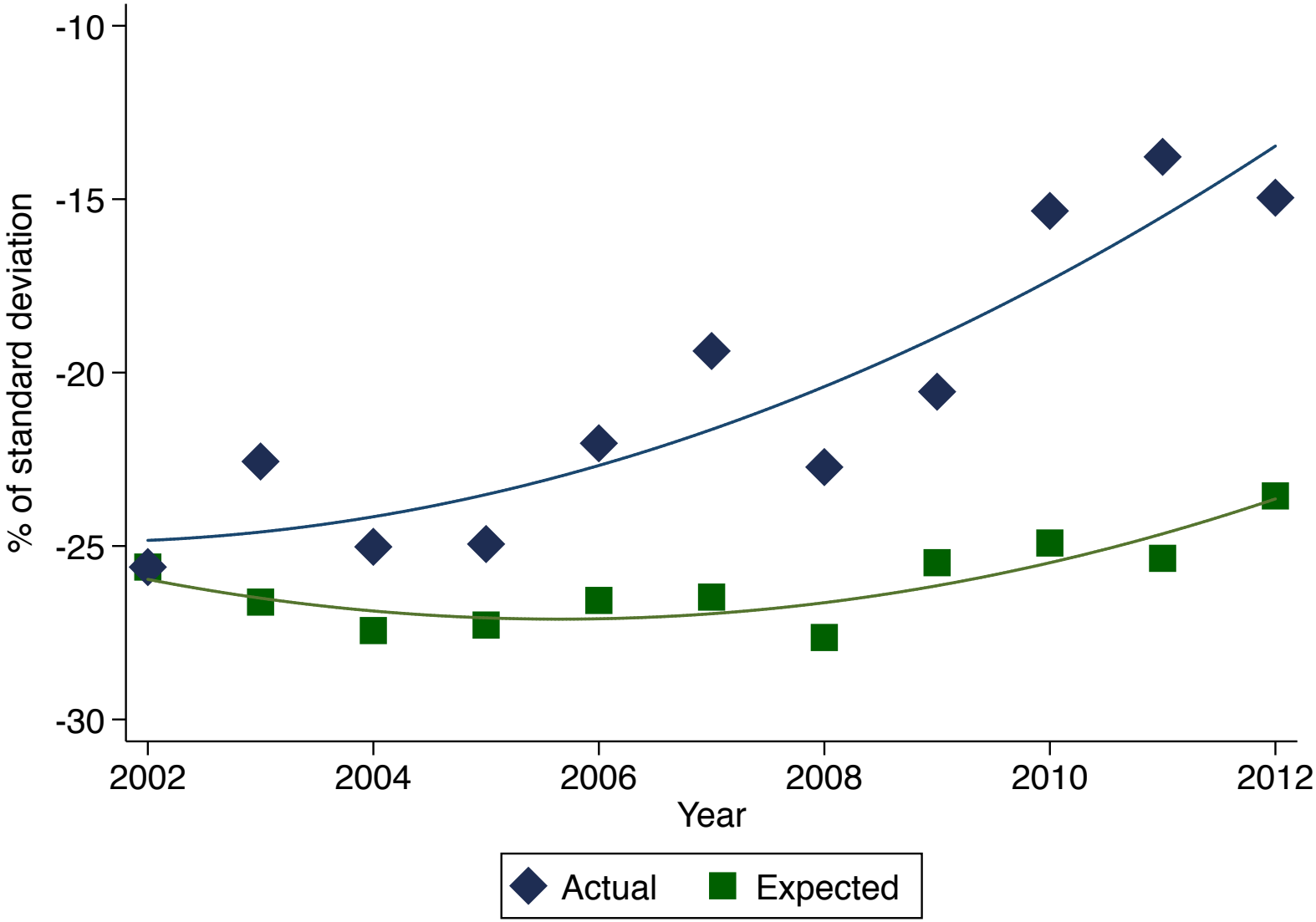
Aggregate Trends in Test Scores



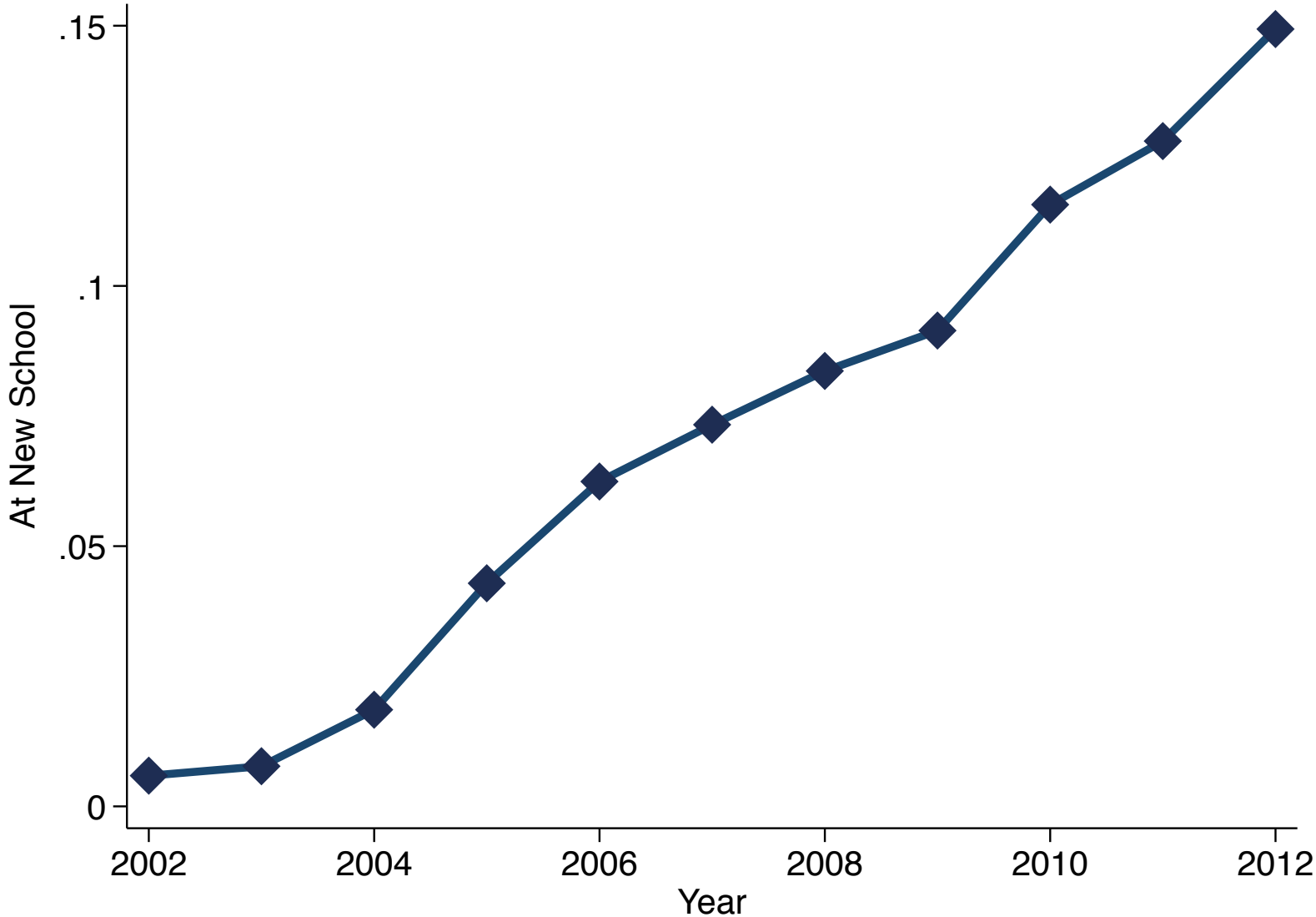
Gap between LA and CA students large, but declining



Gap between LA and CA students large, but declining



By 2012, many students attending newly built schools



Section 3

Empirical Framework

Identifying Facilities Effects

Estimate effects using an event study / DiD framework.

Intuition: students in the same grade and cohort who switch to new schools at different times (or never switch to a new school) form useful counterfactual. Control for:

- Year and grade effects
- Time-invariant individual differences (observed and unobserved)

Causal interpretation relies on assumption that timing of switch as good as random (conditionally)

$\Rightarrow$ Selection problems would have to be *time-varying* and *unobserved*

$\Rightarrow$ **Key feature**: can examine pre-outcomes as placebo test

Estimating Equation

Non-parametric model:

$$y_{igt} = \alpha_i + \gamma_t + \delta_g + \sum_{k=\underline{K}}^{\overline{K}} \beta_k \mathbb{1}(t = t_i^* + k) + \epsilon_{igt}$$

Parametric model:

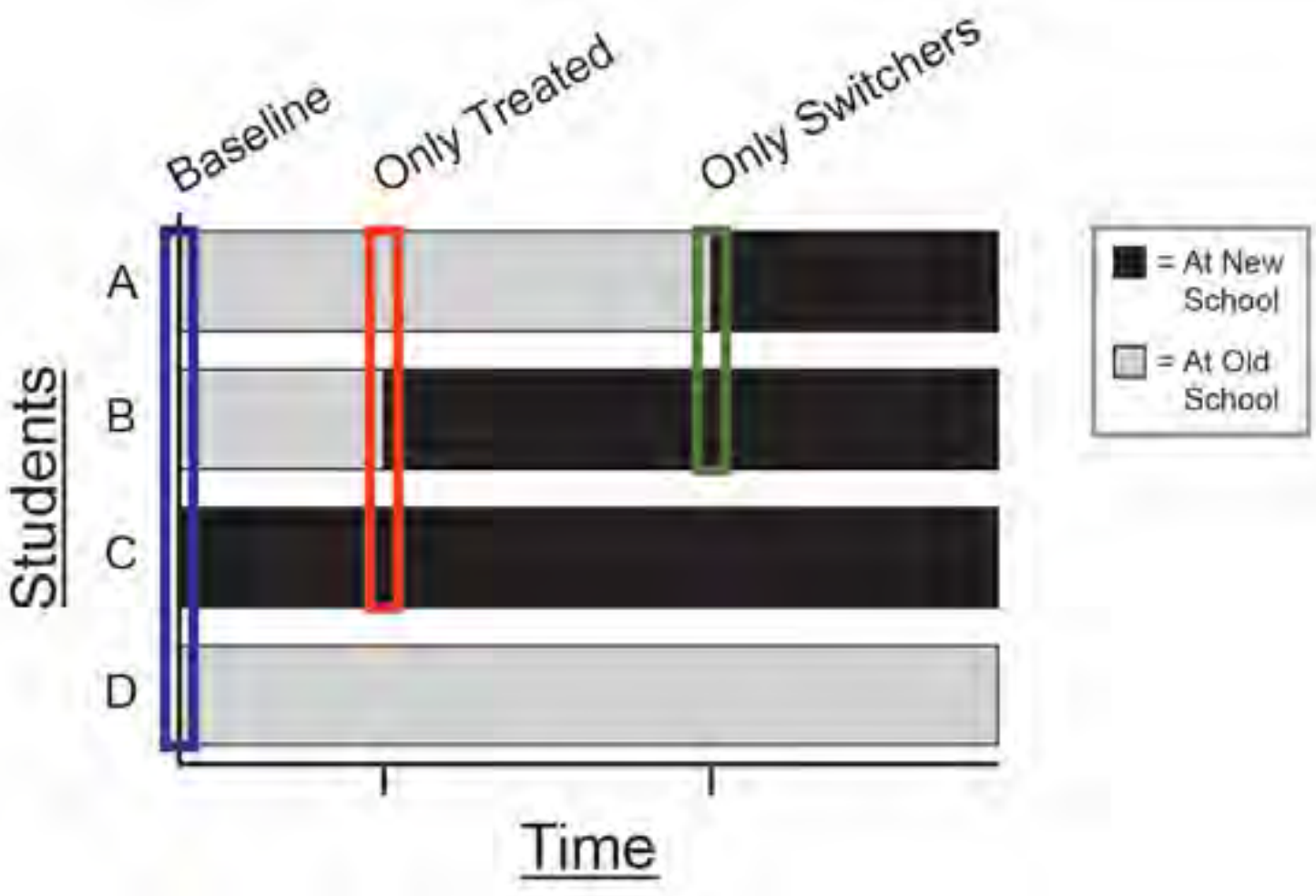
$$y_{igt} = \alpha_g + \alpha_t + \alpha_i + \beta_1 \mathbb{1}(t \geq t_i^*) + \beta_2 \mathbb{1}(t \geq t_i^*) * \tilde{t} + \beta_3 \tilde{t} + \epsilon_{igt}$$

For individual i, grade g, at time t, where:

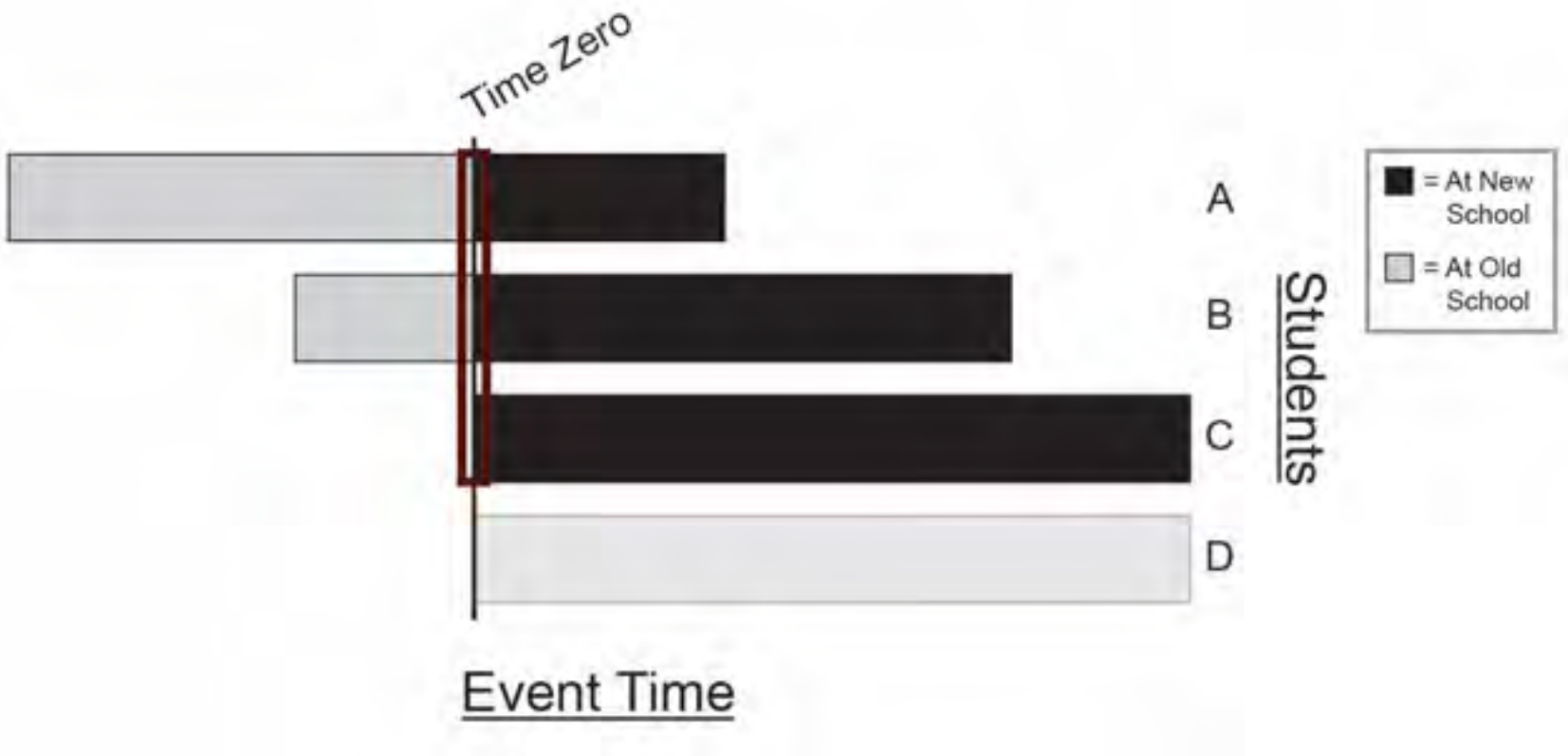
- y_{igt} is student i's outcome
- t_i^* is student's first year in new school
- $\tilde{t}$ is a linear time trend

In non-parametric specifications, bin endpoints at $\underline{K} = -3$ and $\underline{K} = 3$. Standard errors two-way clustered by student and school.

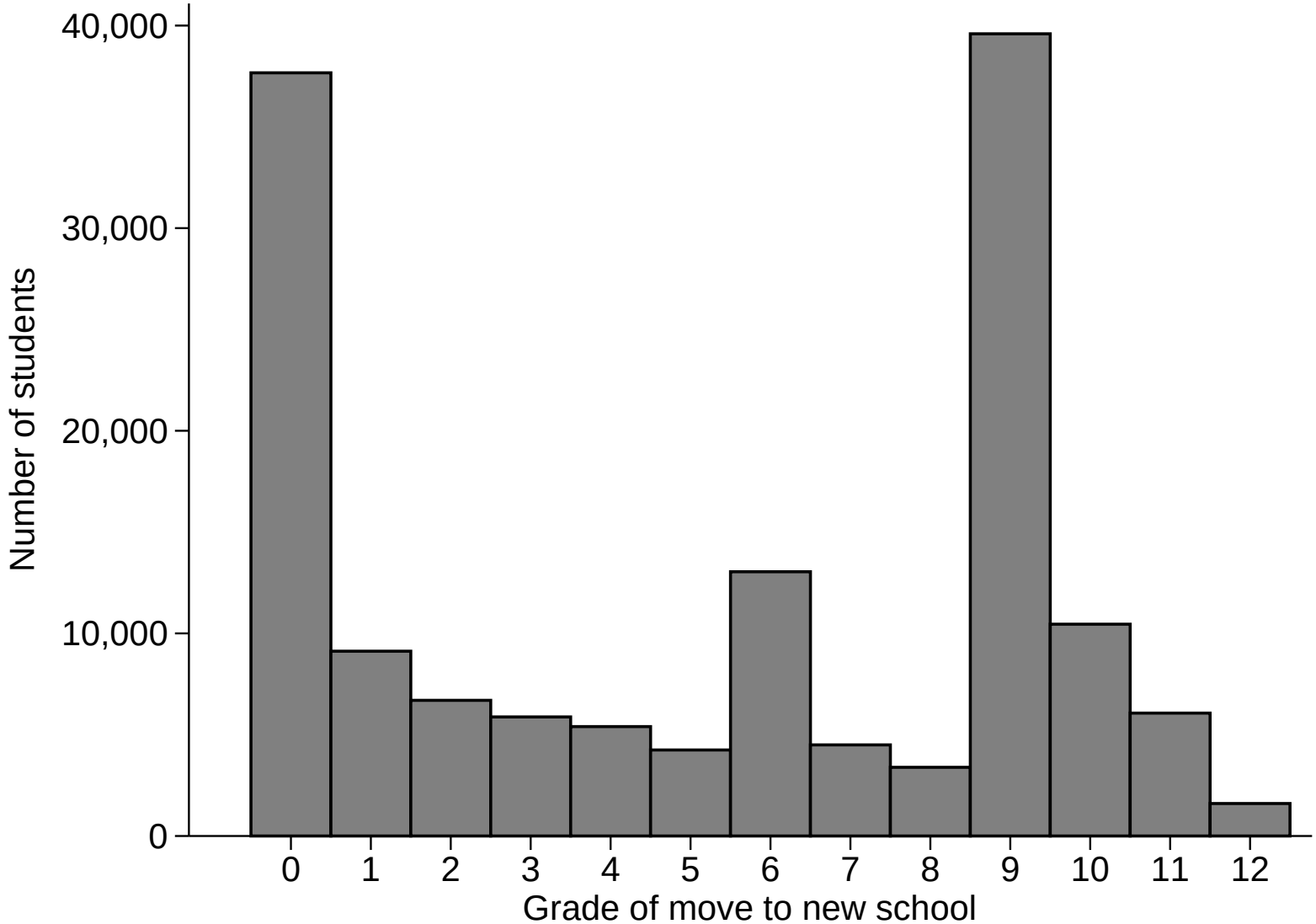
Estimation Strategy



Estimation Strategy



Grade of switch to new school



Balance Table

Table: Balance by treatment group

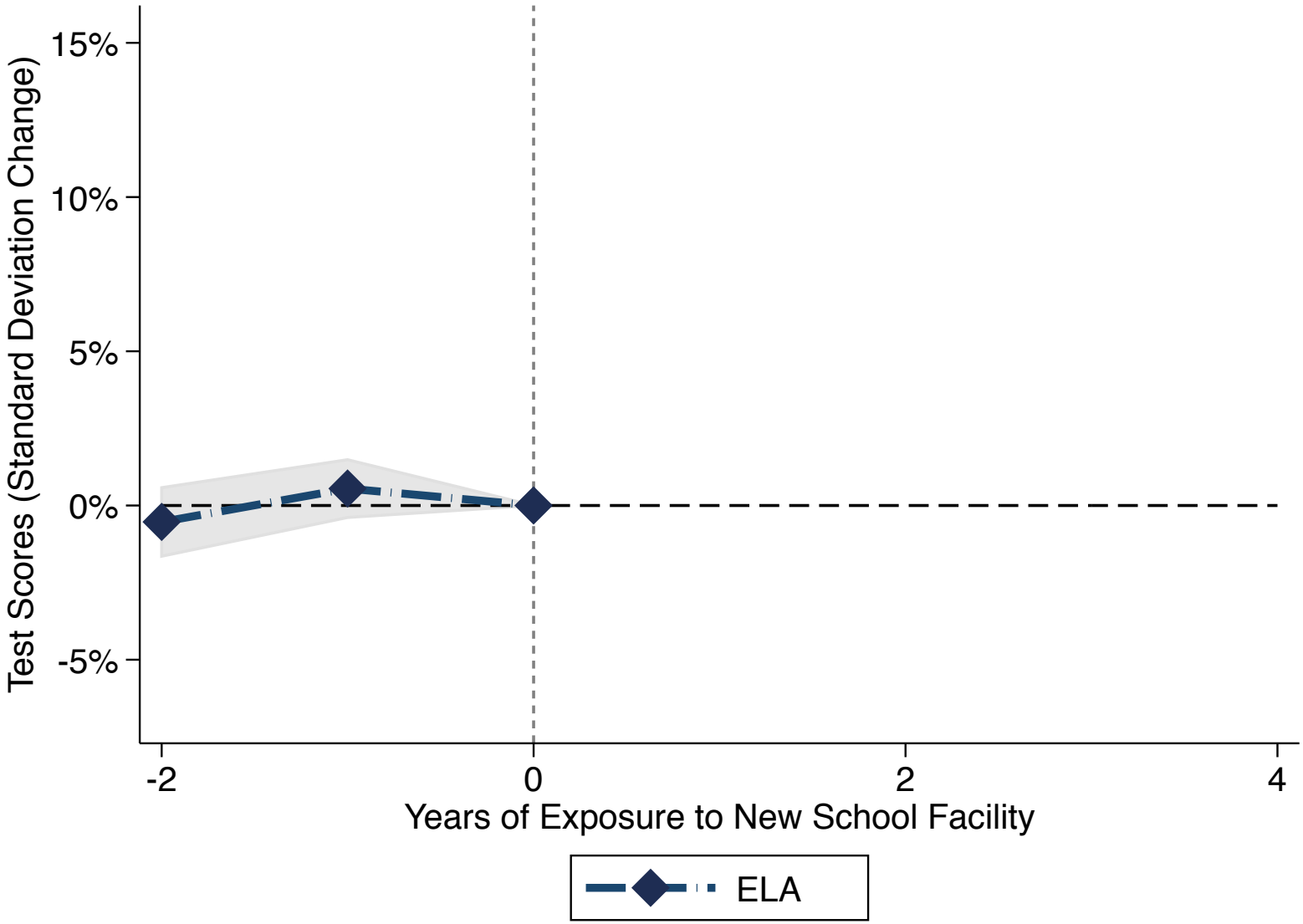
	Never Treated	Always Treated	Switchers	Stayers
Free/reduced lunch	0.75	0.86	0.87	0.85
Parent any college	0.27	0.23	0.16	0.19
Hispanic	0.72	0.85	0.89	0.86
Black	0.11	0.05	0.06	0.05
White	0.10	0.03	0.02	0.04
Asian	0.04	0.04	0.02	0.02
English at home	0.33	0.28	0.17	0.19
Grade	5.7	2.6	5.4	5.6
Math Score ($t = -1$)			-0.35	-0.22
ELA Score ($t = -1$)			-0.52	-0.41
Days Attended ($t = -1$)			156.7	154.7
N	6,711,383	108,749	702,614	1,004,523

Note: Stayers defined as students who have 10% or more of their cohort move to a new school.

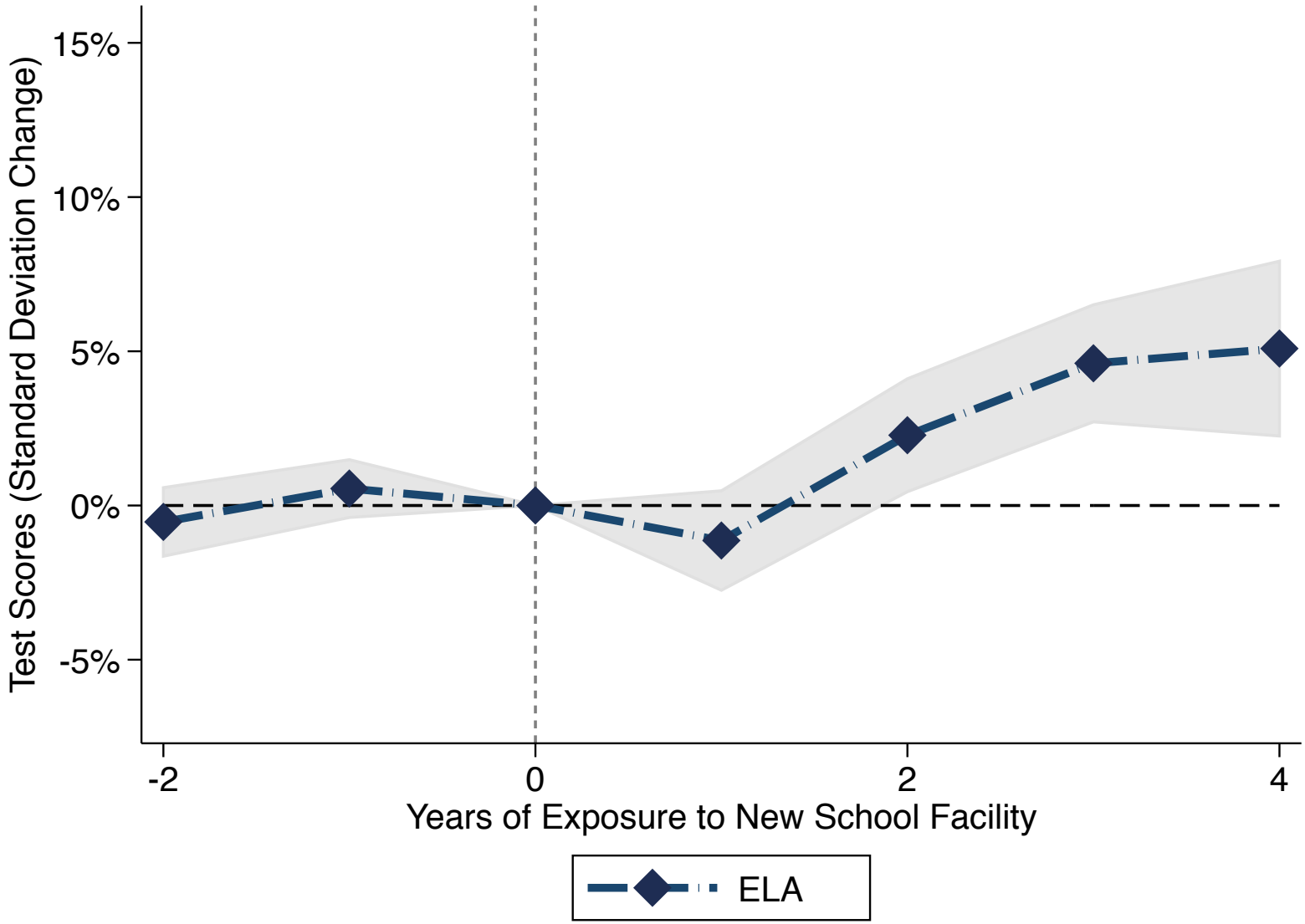
Section 4

Results

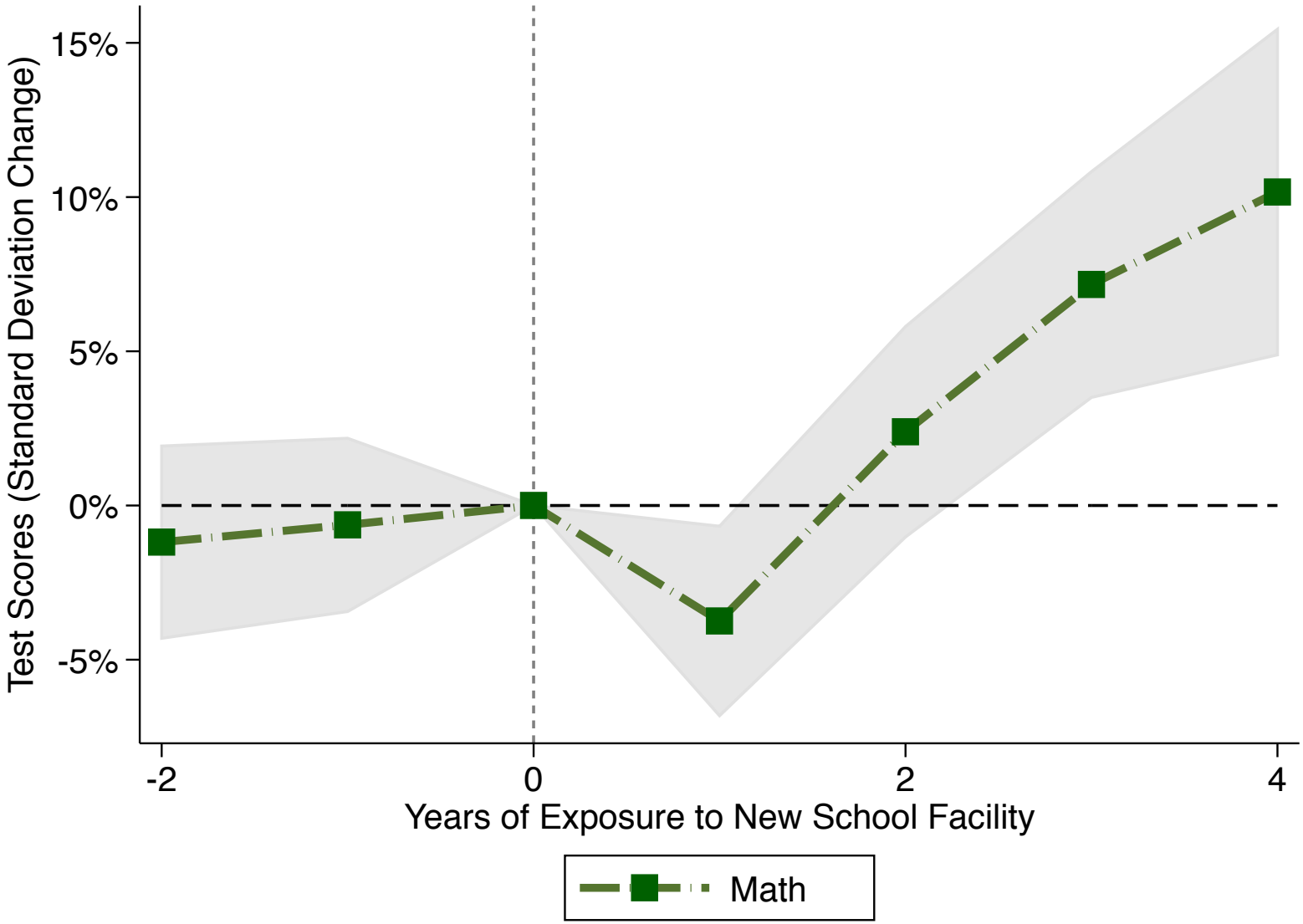
Results: ELA Test Scores (Grades 2-11)



Results: ELA Test Scores (Grades 2-11)



Results: Math Test Scores (Grades 2-7)



Results: ELA Test Scores

Table: DiD Estimates for ELA (Grades 2-11)

	(1)	(2)	(3)	(4)
New School	0.010		-0.004	-0.014
	(0.008)		(0.008)	(0.009)
New School * Trend		0.019***	0.020***	0.017***
		(0.004)	(0.004)	(0.004)
Trend				0.004***
				(0.001)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	4,961,136	4,961,136	4,961,136	4,961,136
N students	1,007,950	1,007,950	1,007,950	1,007,950
N treated students	102,277	102,277	102,277	102,277
N treated schools	132	132	132	132
R2	0.84	0.84	0.84	0.84

Note: OLS regression according to specification (2). Standard errors clustered on students and schools.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

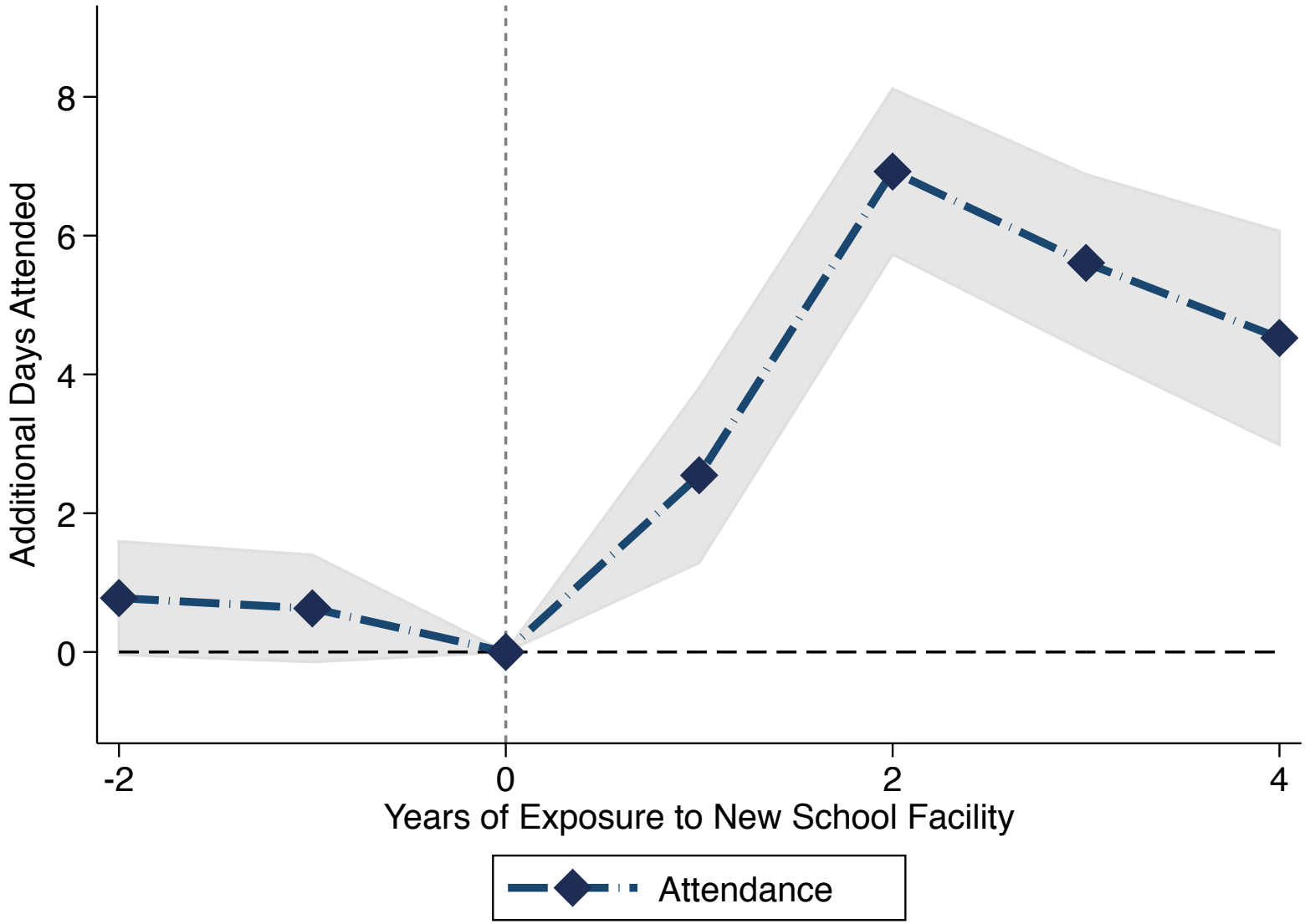
Results: Math Test Scores

Table: DiD Estimates for Math (Grades 2-7)

	(1)	(2)	(3)	(4)
New School	0.004		-0.021	-0.026
	(0.015)		(0.017)	(0.017)
New School * Trend		0.028***	0.033***	0.031***
		(0.007)	(0.008)	(0.008)
Trend				0.004
				(0.002)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	3,095,724	3,095,724	3,095,724	3,095,724
N students	769,827	769,827	769,827	769,827
N treated students	89,439	89,439	89,439	89,439
N treated schools	82	82	82	82
R2	0.82	0.82	0.82	0.82

Note: OLS regression according to specification (2). Standard errors clustered on students and schools.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results: Attendance (Grades K-12)



Results: Attendance

Table: DiD Estimates for Days Attended

	(1)	(2)	(3)	(4)
New School	3.878***		3.269***	3.294***
	(0.544)		(0.590)	(0.630)
New School * Trend		1.707***	0.821***	0.830***
		(0.213)	(0.191)	(0.199)
Trend				-0.010
				(0.084)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	5,615,447	5,615,447	5,615,447	5,615,447
N students	1,163,271	1,163,271	1,163,271	1,163,271
N treated students	127,940	127,940	127,940	127,940
N treated schools	150	150	150	150
R2	0.51	0.51	0.51	0.51

Note: OLS regression according to specification (2). Standard errors clustered on students and schools.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

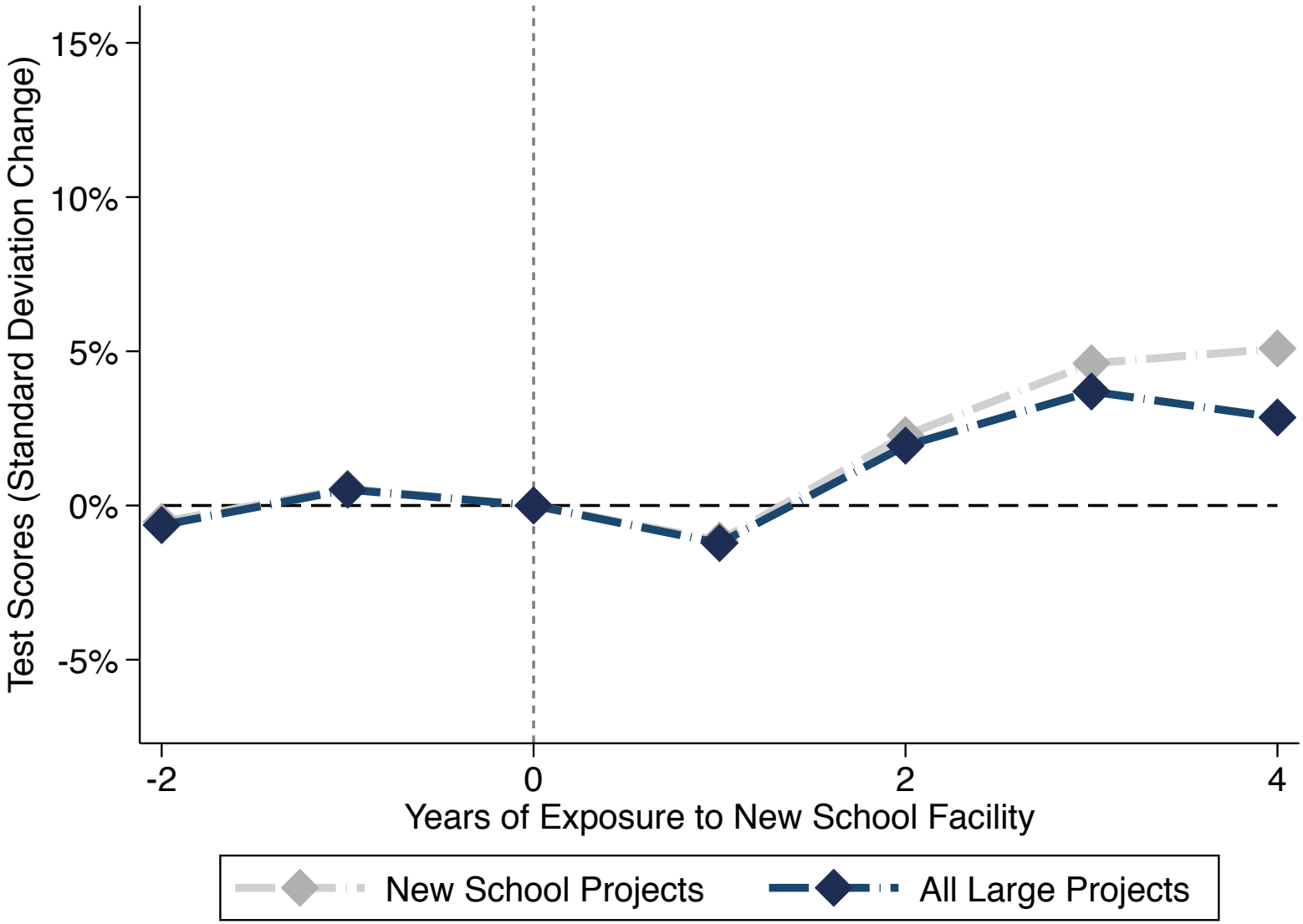
Results: Robustness

Table: DiD Estimates: Robustness

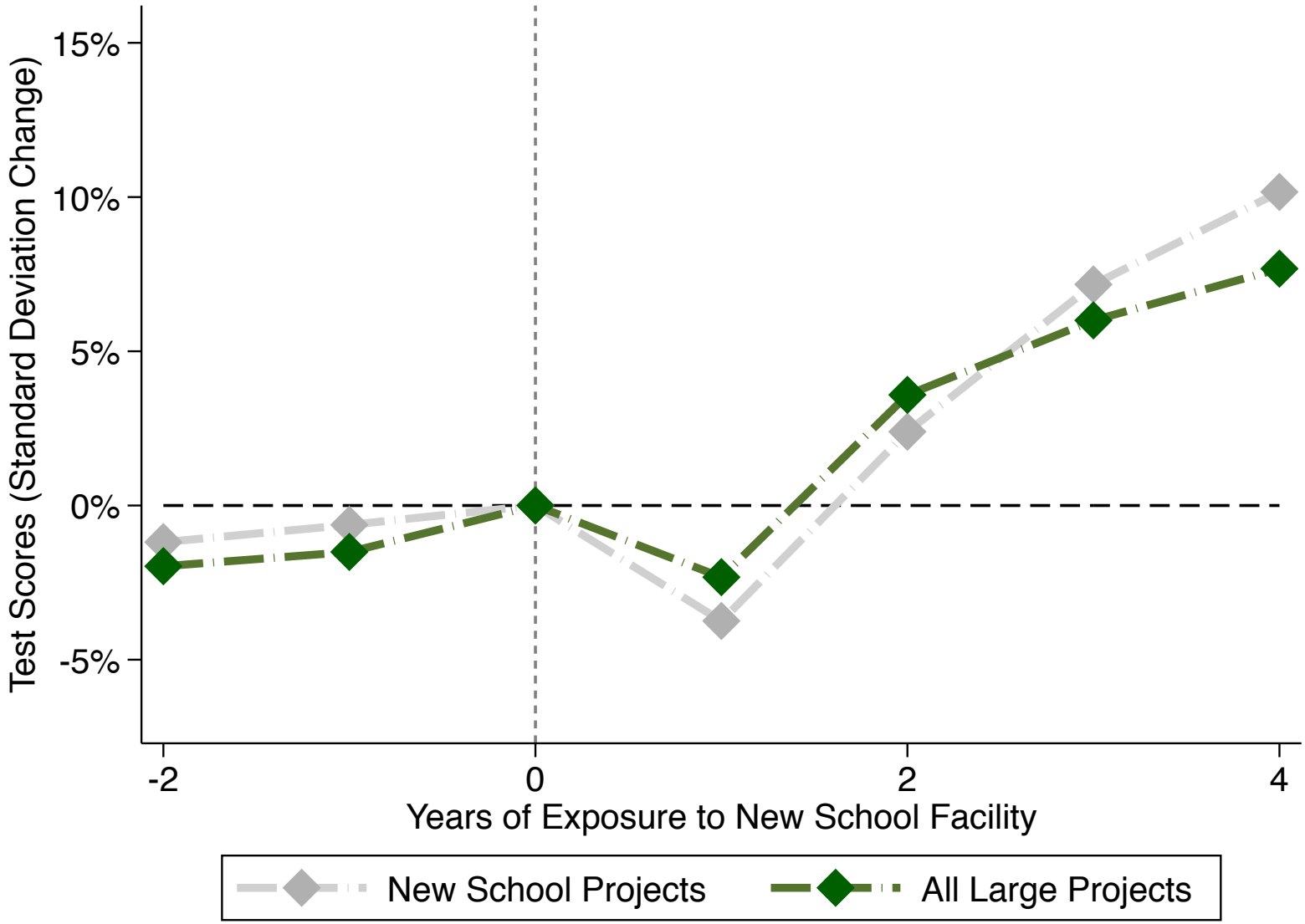
	Baseline	Only Treated	Only Switchers	Balanced
ELA Score				
New School * Trend	0.019***	0.018***	0.016***	0.027*
	(0.004)	(0.005)	(0.005)	(0.014)
Math Score				
New School * Trend	0.028***	0.034***	0.035***	0.059*
	(0.007)	(0.011)	(0.012)	(0.033)
Days Attended				
New School	3.878***	3.869***	4.294***	8.521***
	(0.544)	(0.779)	(0.790)	(1.654)

Note: OLS regression according to specifications (1) (row 3) and (2) (rows 1 and 2). Standard errors clustered on students and schools.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

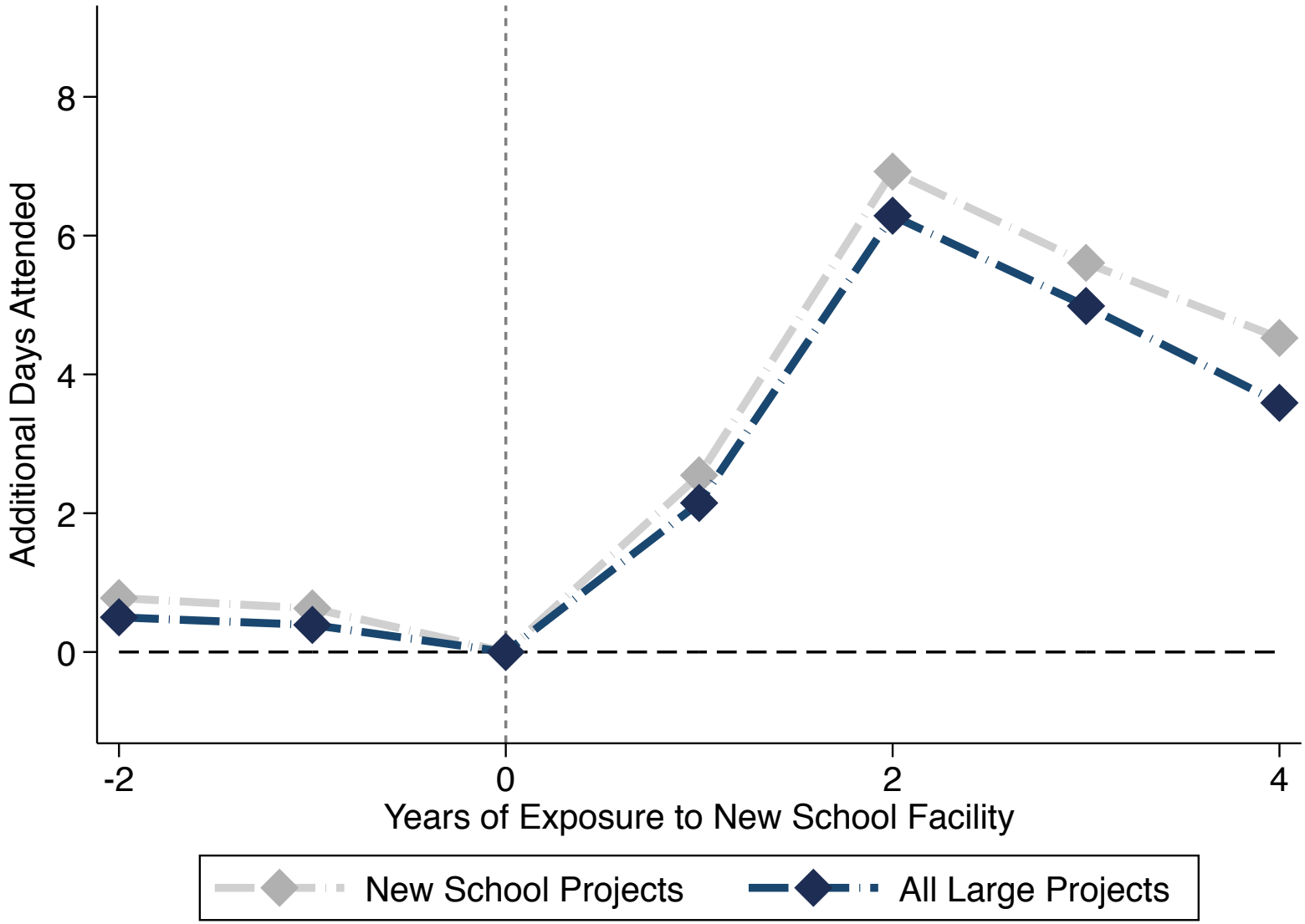
Comparing to All Major Projects: ELA



Comparing to All Major Projects: Math



Comparing to All Major Projects: Attendance



Interpreting Size of Impacts

1. Significant reduction of **gap to California** average:
   - ELA gap of 28% of SD and math gap of 23%

   $\Rightarrow$ 5% ELA score increase: 18% of gap closed

   $\Rightarrow$ 10% Math score increase: 45% of gap closed
2. **Equivalent** to meaningful **class size reduction** effects:
   - 1/3 class size reduction leads to 18% increase

   $\Rightarrow$ ELA score increase: 10% reduction in class size

   $\Rightarrow$ Math score increase: 20% reduction in class size
3. **Equivalent** to a large **increase in instructional days**:
   - 25% increase is a 35 in math, 45 in ELA

   $\Rightarrow$ ELA score incrase: 9 more instructional days

   $\Rightarrow$ Math score increase: 14 more instructional days

Wider Implications for the District

1. **Adult earnings**:
   - \$3,530 for 10% SD increase in test scores (Chetty et al. 2011)
   - About 442,274 student-years treated

   $\Rightarrow$ \$225 million up to 2012

   $\Rightarrow$ \$740 million up to 2022 with capped occupancy
2. **LAUSD operational funds**:
   - About 3.5 additional ADA and \$42 per student-year ADA

   $\Rightarrow$ ADA increase of 1,547,959

   $\Rightarrow$ \$65 million in increased operational funds

This does not take into account district-wide effects!

Section 5

Mechanisms

Unpacking the "Black Box"

Why do students benefit from attending new schools?

Non-exclusive list of potential reasons:

1. Facility quality and overcrowding
2. Switch from multi- to single-track
3. Staff and teacher sorting
4. Student health/motivation
5. Teacher health/motivation
6. Way-to-school quality

Unpacking the "Black Box"

Why do students benefit from attending new schools?

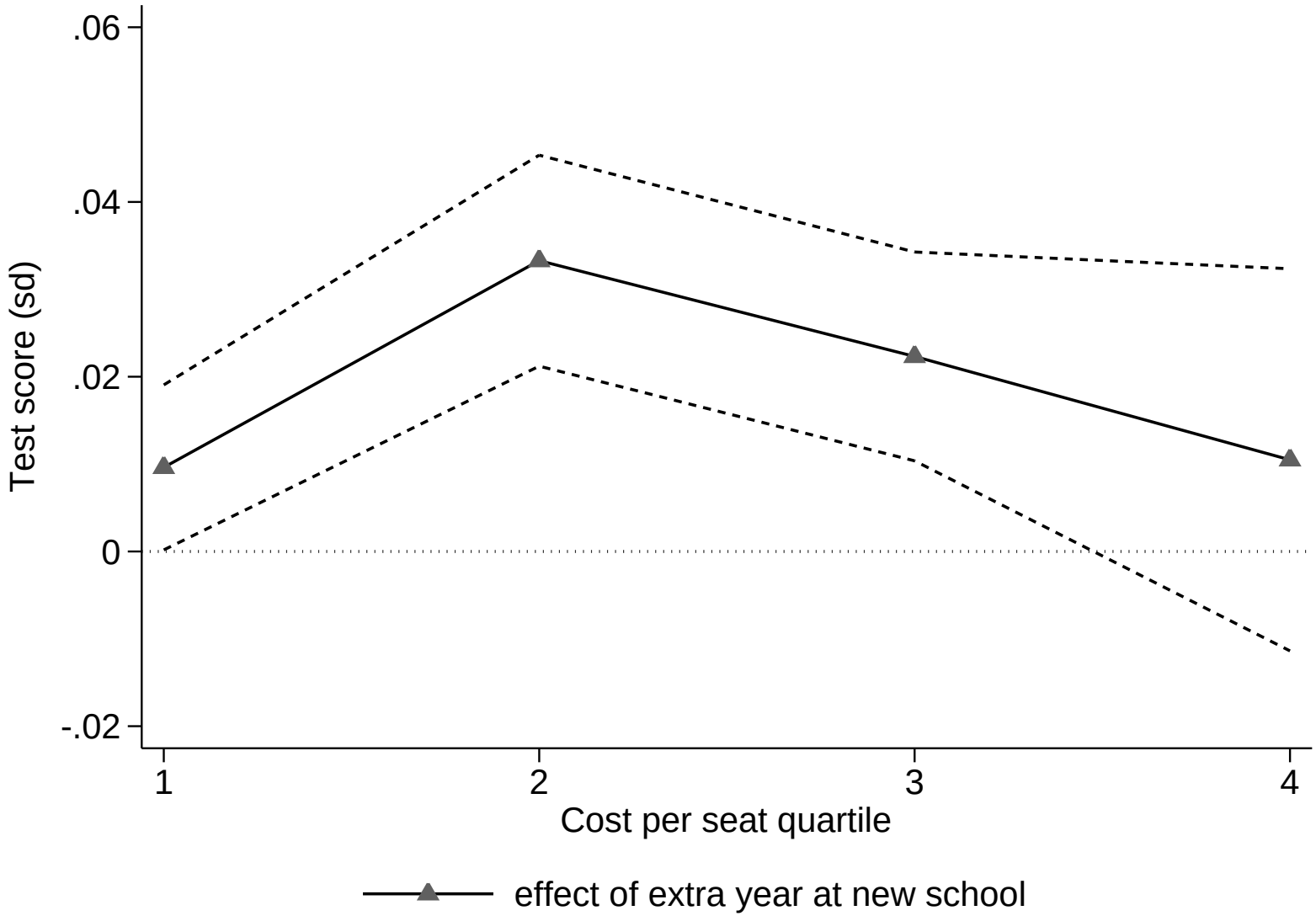
First pass: add interactions to baseline parametric model*
($\mathbb{1}(t \geq t_i^*) * \tilde{t}$ or $\mathbb{1}(t \geq t_i^*)$) to examine heterogeneity by

- Cost per pupil of new construction
- Prior achievement quartiles
- Prior school condition / prior school congestion quartiles
- Prior school calendar (multi vs single)
- Grade & parental education / SES

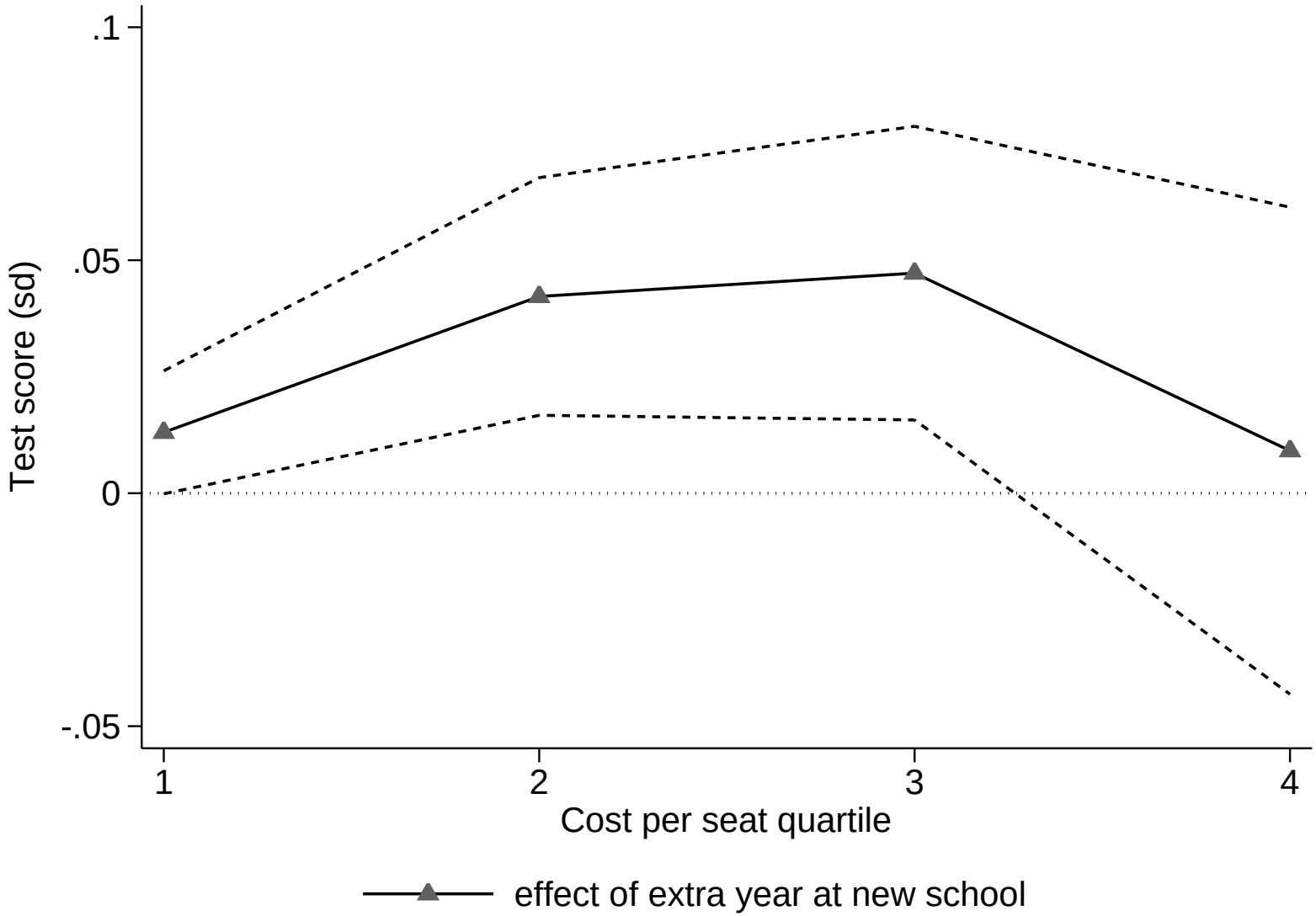
*_Recall_:

$$y_{igt} = \alpha_g + \alpha_t + \alpha_i + \beta_1 \mathbb{1}(t \geq t_i^*) + \beta_2 \mathbb{1}(t \geq t_i^*) * \tilde{t} + \beta_3 \tilde{t} + \epsilon_{igt}$$

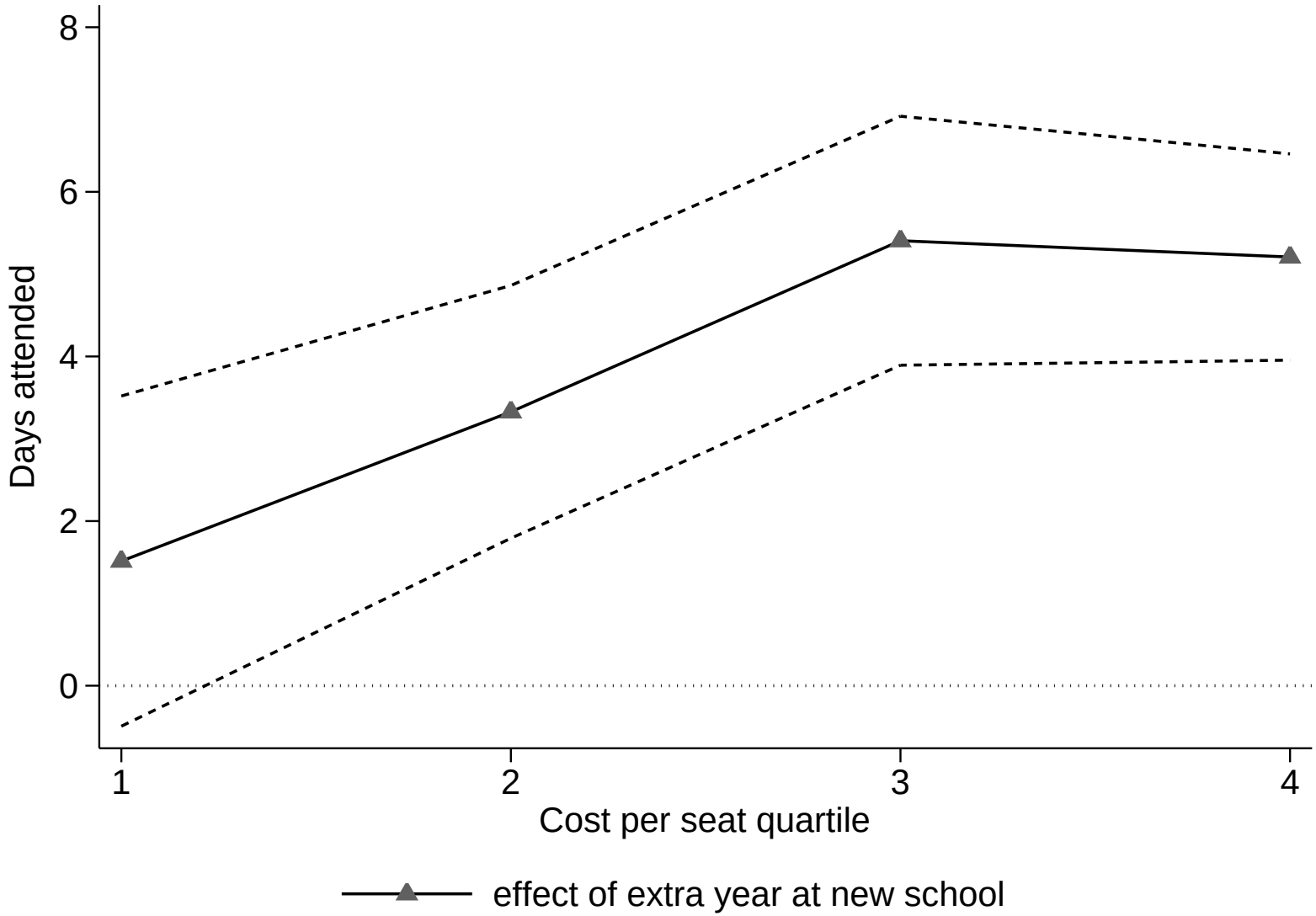
Heterogeneity by Cost per pupil: ELA



Heterogeneity by Cost per pupil: Math



Heterogeneity by Cost per pupil: Attendance



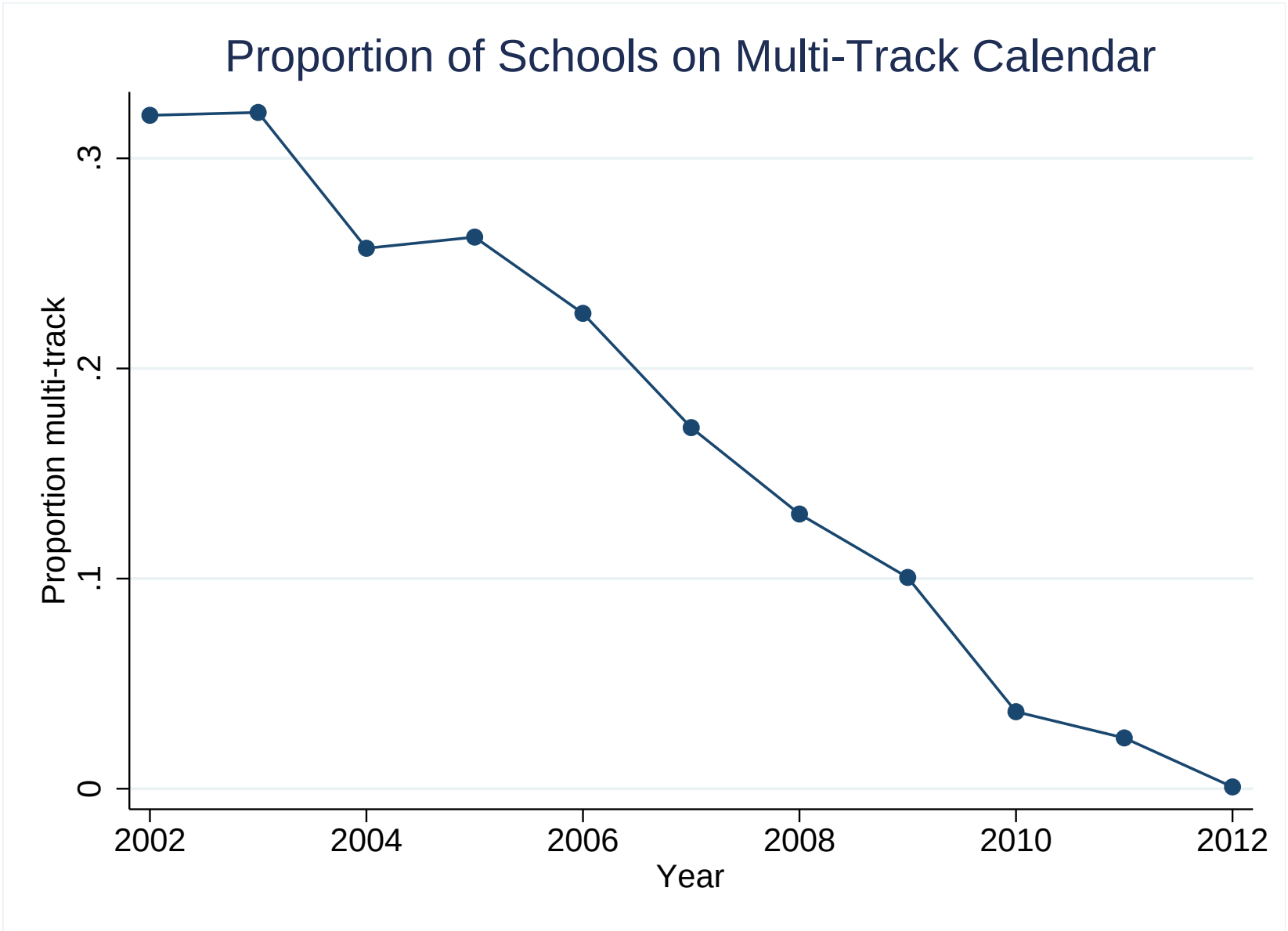
Heterogeneity: Other Results

Using same methodology, little sign of heterogeneity by:

- Prior school condition (roughly defined)
- Prior school congestion (roughly defined)
- Parental background / SES

By grade: Test score effects slightly larger in elementary school; attendance effects larger in later grades (middle and high school).

Unpacking the "Black Box": Multi-Track Calendars?



Unpacking the "Black Box": Multi-Track Calendars?

Decomposing results by prior "track", find that:

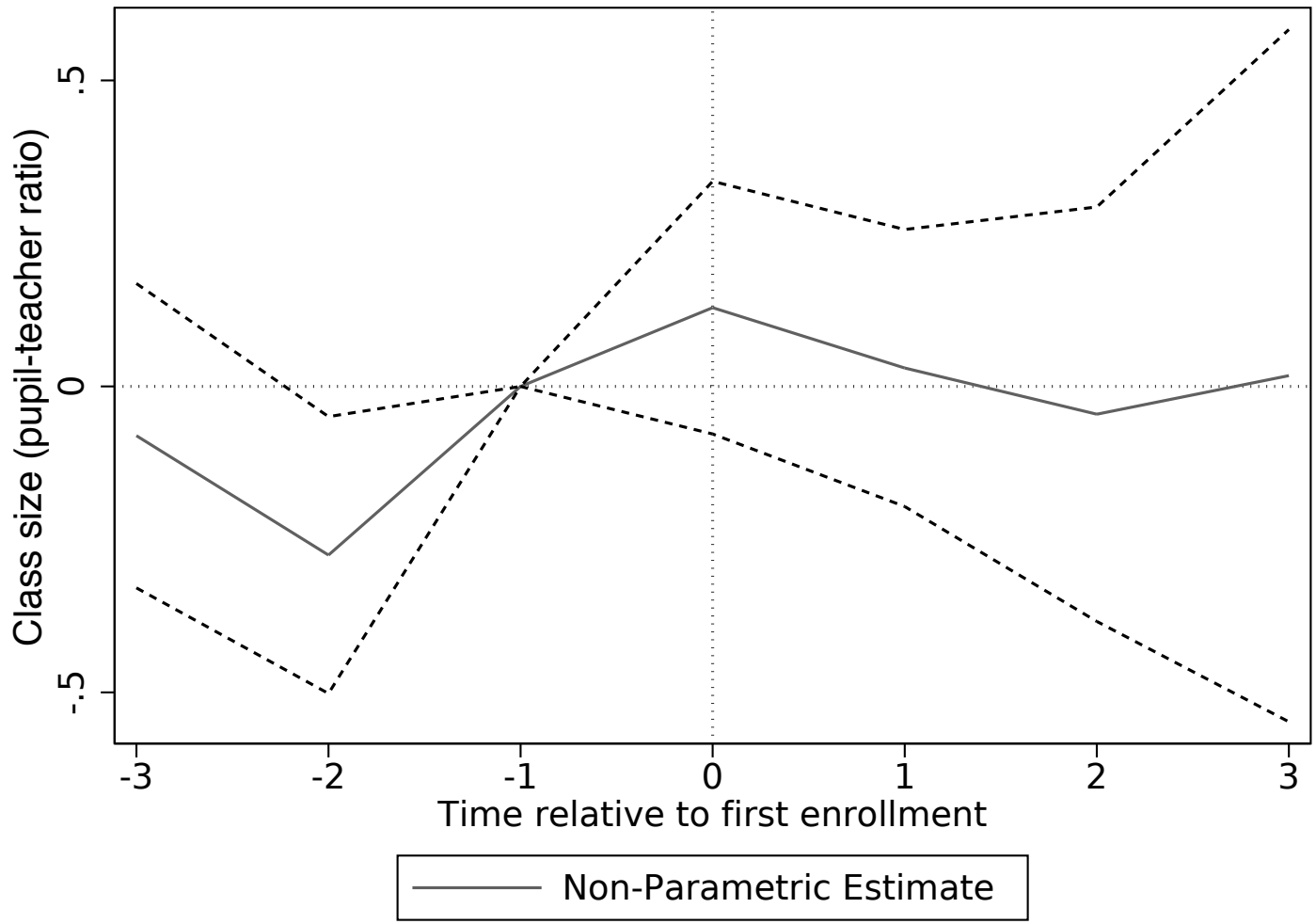
- Attendance effects 2x larger when going from multi- to single-track
  - $\Rightarrow$ Mechanical; more instructional days at single track
- Insignificant differences in ELA and math effects

Eliminating multi-track calendars may still have had positive impacts on *district-wide* student outcomes, but we estimate little difference in *relative* outcomes of students at new constructions.

$\rightarrow$ "Stayers" still saw reductions in overcrowding / multi-track scheduling

Unpacking the "Black Box": Class Size?

Can estimate analogous event study with class size as student outcome, y_{igt}:



Unpacking the "Black Box": Teachers

Effects of new facilities could be mediated through teachers.

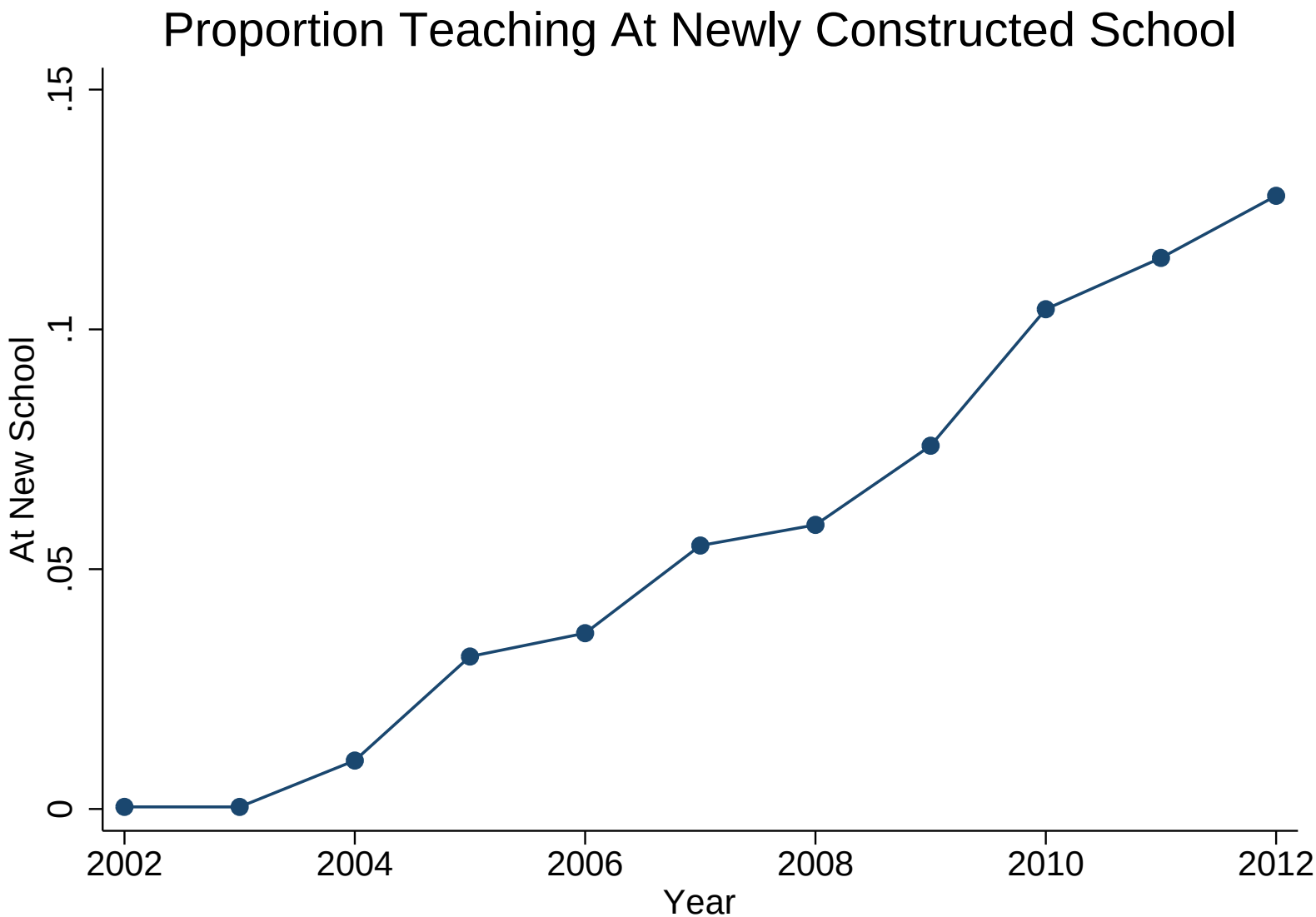
Better/new facilities might:

1. Attract better teachers from *within* district
2. Attract better teachers from *outside* district
3. Improve teacher productivity

Today, try to address **(1)** and **(2)** using teacher observables (age, education, experience).

Eventually, use student outcome data to quantitatively assess **(3)**.

Teachers at New Schools

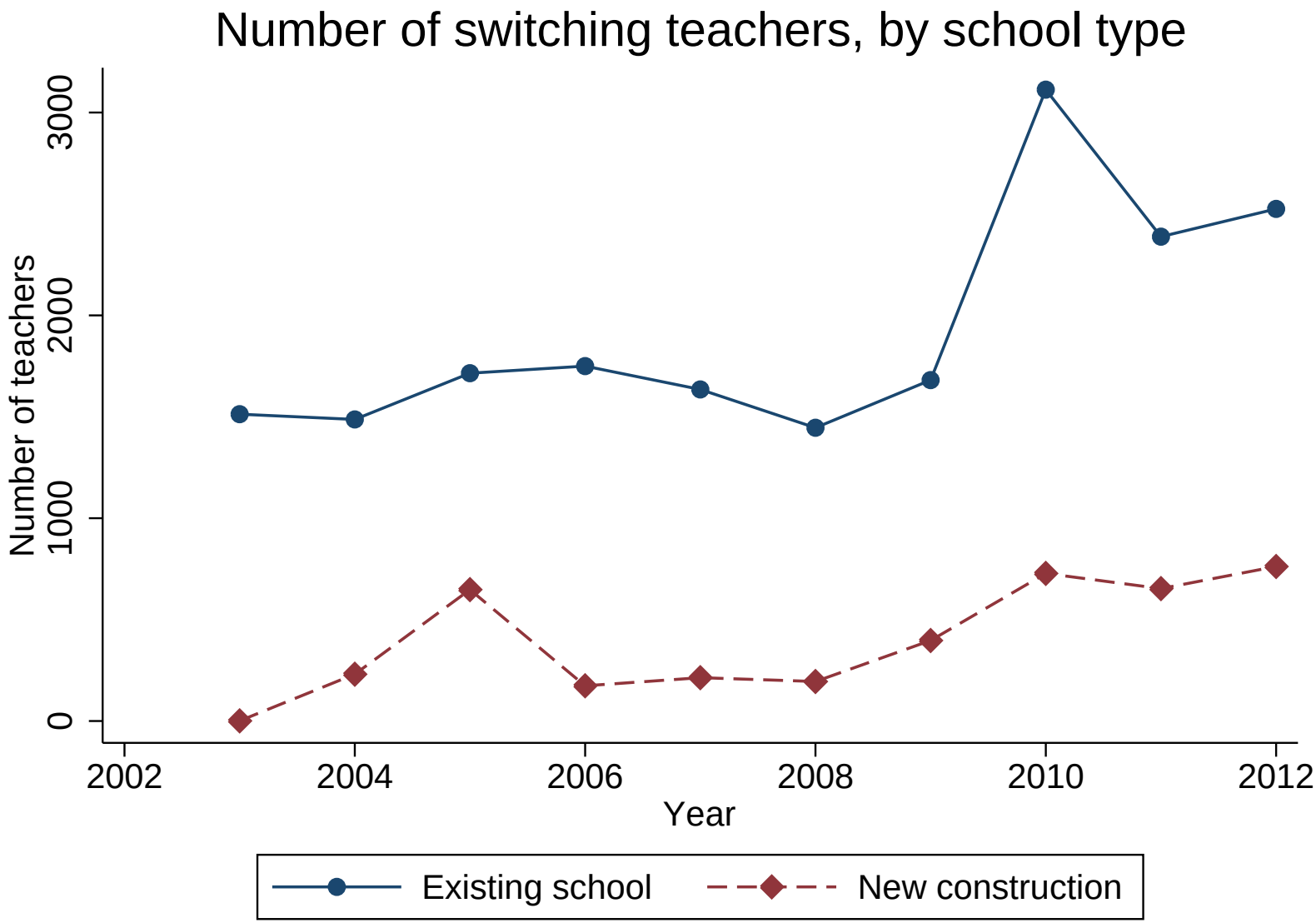


Teacher Selection into New Schools

Table: Teacher Observables, by School Type

	Existing school		Newly built school	
	Mean	Median	Mean	Median
Experience: in district	12.41	10	10.07	9
Experience: overall	13.07	11	10.29	9
Has MA+ degree	0.37	0	0.40	0
Age	43.88	43	40.30	38
Female	0.69	1	0.69	1
Num yrs in data	8.39	9	8.36	9
Observations	388289		19920	

Teachers Switching

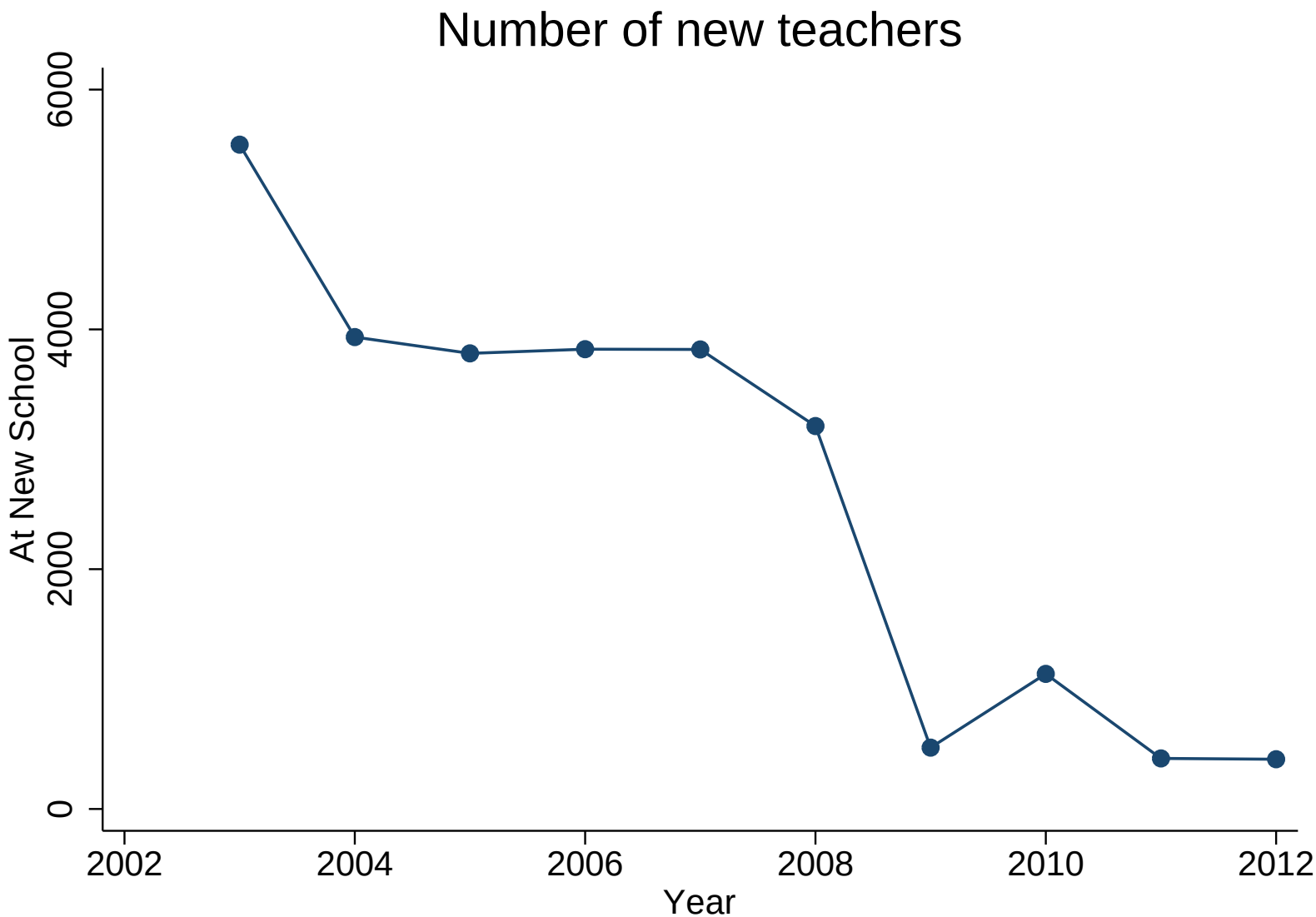


Teacher Selection into New Schools

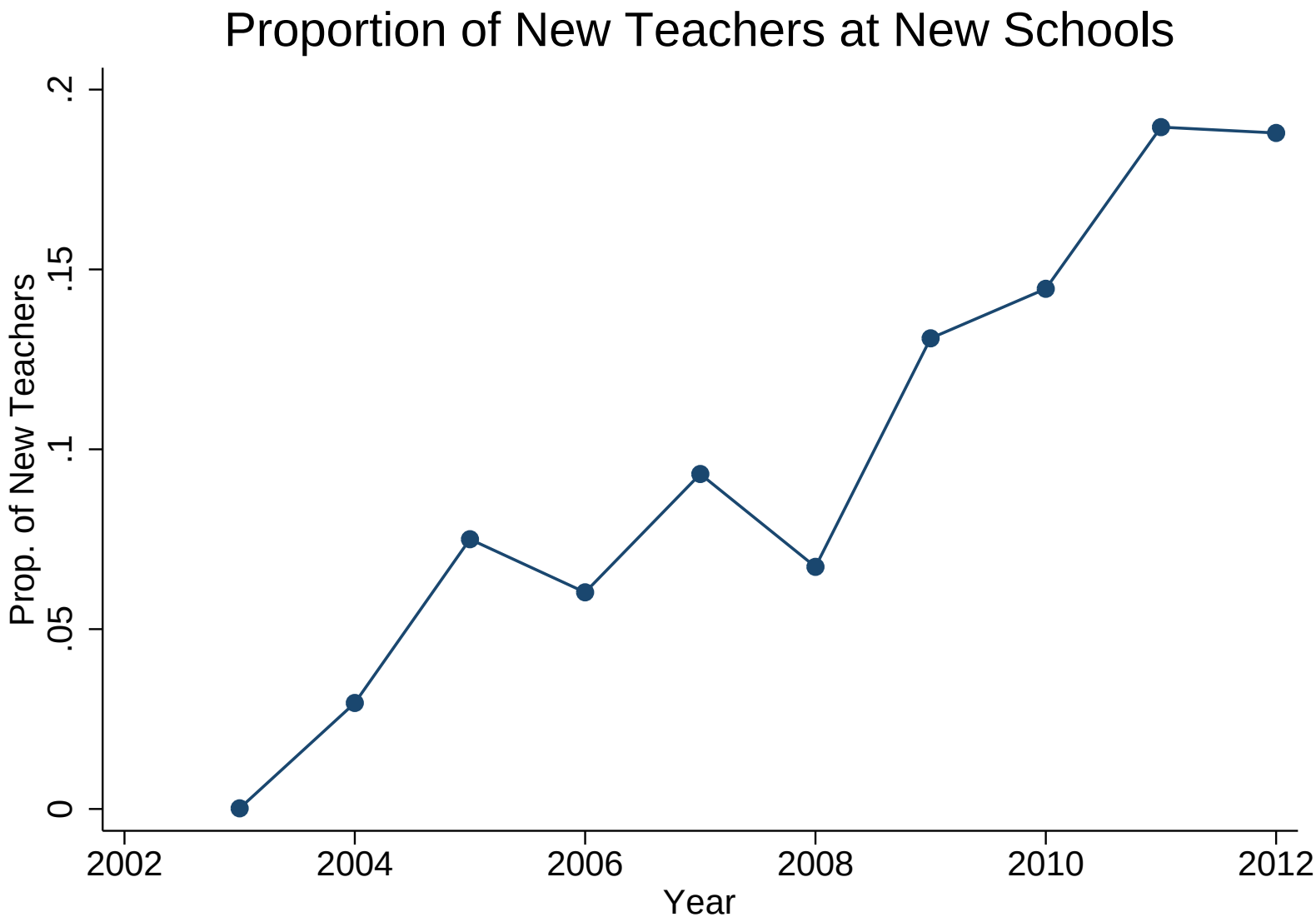
Table: Teacher "Switchers", by School Type

	Existing school		Newly built school	
	Mean	Median	Mean	Median
Experience: in district	11.74	10	10.58	9
Experience: overall	12.19	10	10.85	9
Has MA+ degree	0.48	0	0.41	0
Age	42.90	41	40.80	39
Female	0.66	1	0.68	1
Num yrs in data	8.48	9	8.75	9
Observations	18979		3923	

New Teachers



New Teachers

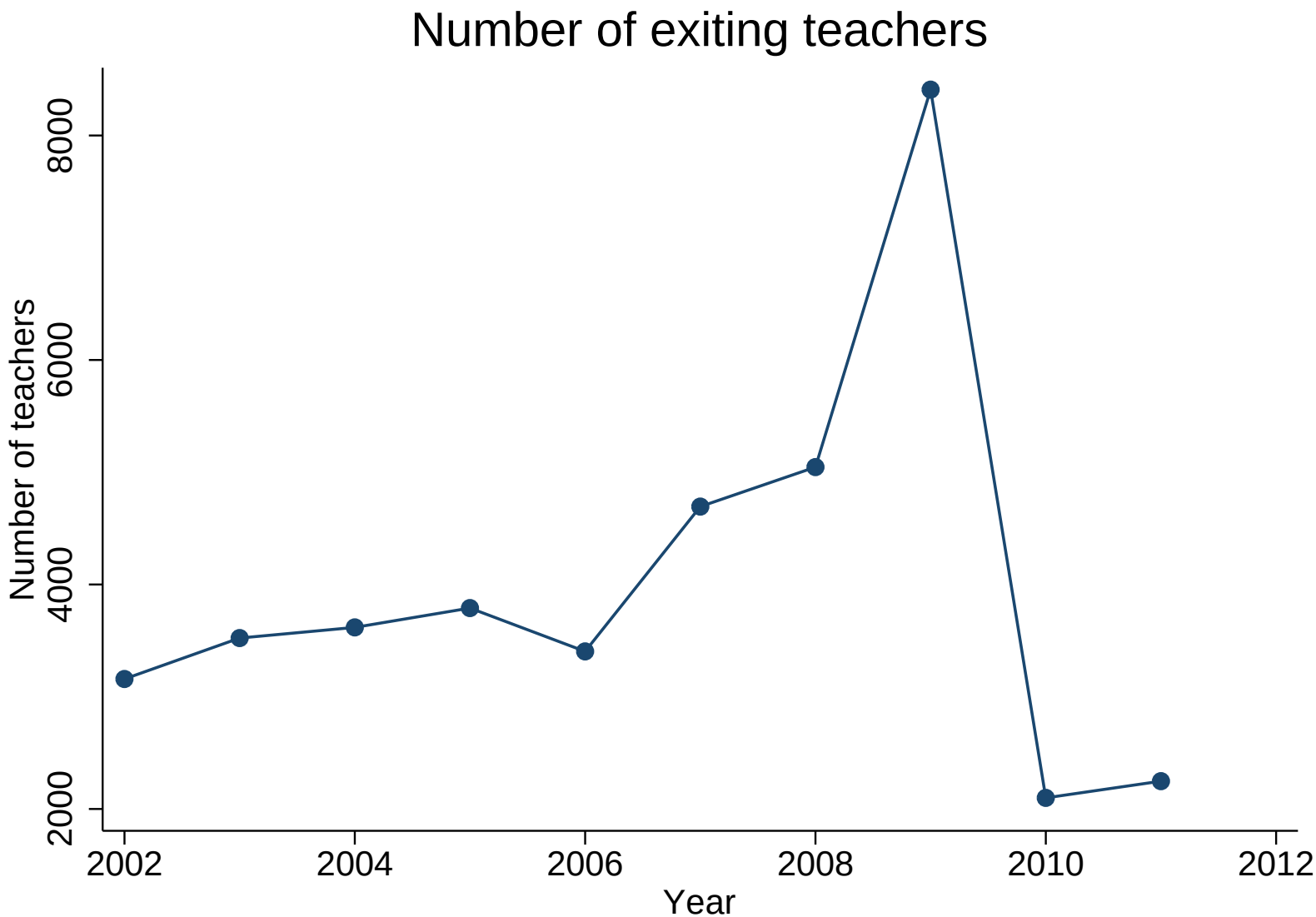


Teacher Selection into New Schools

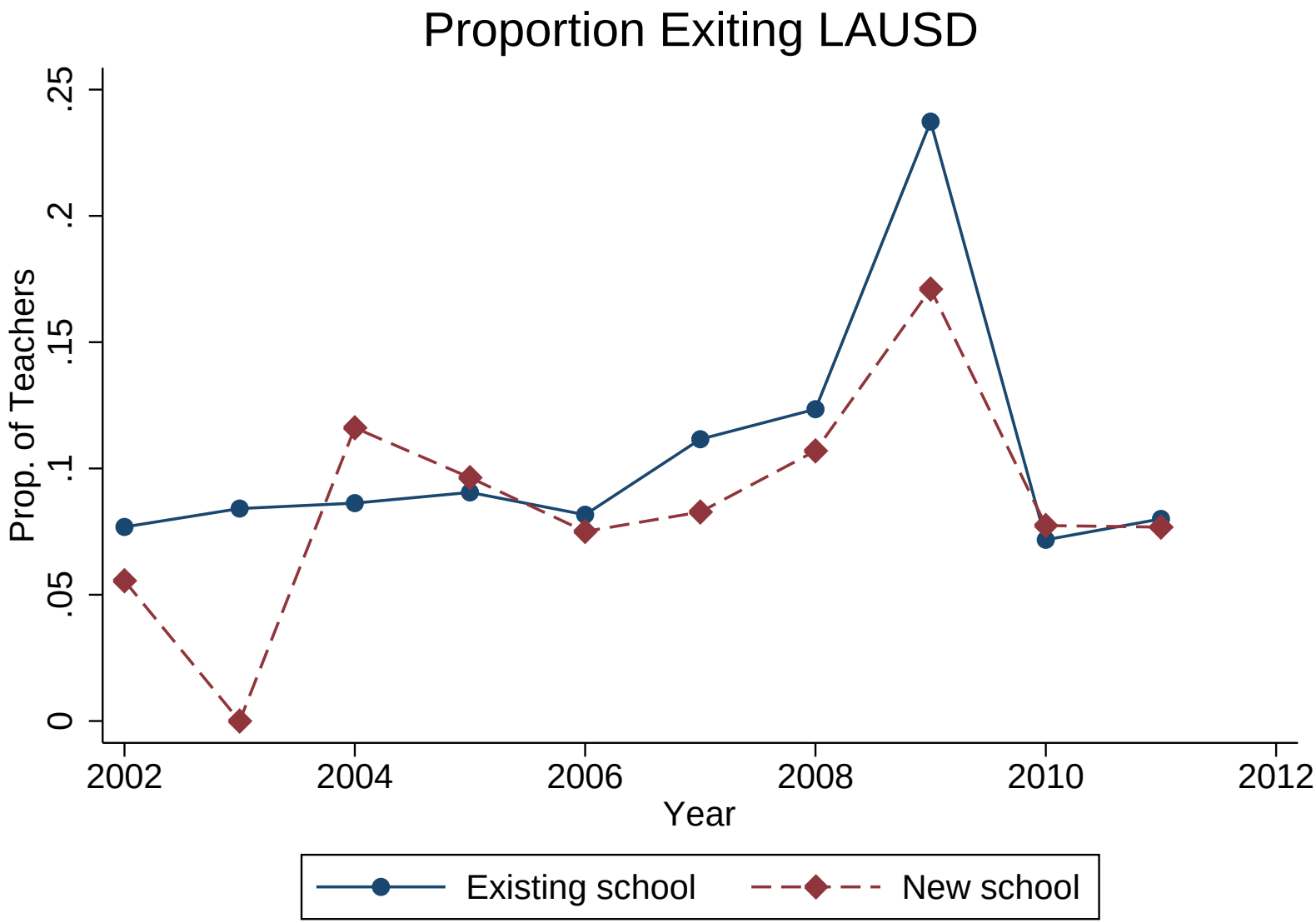
Table: New Teachers, by Starting School Type

	Existing school		Newly built school	
	Mean	Median	Mean	Median
Experience: overall	3.81	1	1.93	1
Has MA+ degree	0.23	0	0.20	0
Age	35.00	31	32.89	29
Female	0.69	1	0.67	1
Num yrs in data	4.38	4	4.14	4
Observations	22840		1254	

Teacher Persistence



Teacher Persistence



Unpacking the "Black Box": Teachers

If anything, teachers at new schools are (slightly) negatively selected on observables:

1. Less experience overall, and within LAUSD
2. Switching teachers slightly less experienced, educated
3. Higher share of new teachers at new schools
4. New teachers begin with slightly less experience, education

But, magnitude of differences small.

$\Rightarrow$ Observable differences unlikely to explain estimated facilities impacts

$\Rightarrow$ Unobserved differences? Teacher value-added?

Unpacking the "Black Box": Takeaways

Find little evidence of strong mediating effects of:

1. Class size
2. Observed teacher differences
3. Switch away from multi-track calendars
4. Physical congestion (roughly defined)
5. Prior facility condition (roughly defined)

Importantly, **(4)** and **(5)** could still be crucial factors, but hard to examine by only looking at *heterogeneity between students* switching from more or less congested/deteriorated schools.

Conclusions

Each additional year of attending a newly constructed school is associated with robust gains in test scores and attendance:

$\Rightarrow$ 1.9% of SD increase in ELA test scores

$\Rightarrow$ 2.8% of SD increase in math test scores

$\Rightarrow$ 4 additional days in attendance

Results imply large reductions in achievement gap relative to average CA student.

Can rule out non-facilities mediators, including class size, school calendar type, and teacher observables.

Further work necessary to understand:

1. Impacts of new facilities spending on teacher productivity and recruitment.
2. Cost/benefit of new facilities spending (vs. instructional spending, etc).