The Impact of Schooling Intensity on Student Learning: Evidence from a Quasi-Experiment

Vincenzo Andrietti, Università "d'Annunzio"

Xuejuan Su, University of Alberta

Duration and **intensity** of schooling:

- Jointly shape the **curriculum** and determine its **effectiveness**
- **Schooling duration effects** (i.e., years of schooling): more lead to
  - better labor market outcomes (Card, 2001, Heckman, 2006)
  - better learning in next educ. stage (Morin, 2013, Krashinsky, 2013)
- **Schooling intensity effects**: (i.e., curriculum content covered in a year) can be hard to gauge
  - If additional instruction time used to cover same set of topics, **schooling intensity does not change**: students expected to do better even if marginal returns are diminishing
  - If additional instruction time used to cover more academic content **schooling intensity increases**: students may struggle to keep up

Literature

- Large literature analyzing the effects of instruction time on learning outcomes in settings where **schooling intensity does not change**
- For example, Lavy (2015) and Rivkin and Schiman (2015) find that instruction time has significant positive effects on PISA test scores
- **Schooling intensity** as a **determinant of the effectiveness of the schooling process** is an empirically unexplored topic

Major identification challenges

1. **Measurement difficult**: Even if one has data on the amount of instruction time in a school year, it is typically hard to quantify **the amount of academic content covered in those hours**. Same instruction time may correspond to different schooling intensities
2. **Endogeneity** of observed variations in intensity levels:
   - Students self-select into suitable intensity levels to improve outcomes
   - Teachers adjust the intensity level to better serve their students.

   Need for **exogenous** variation in schooling intensity

The G8 reform in Germany as a quasi-natural experiment

- **G8 reform**: **academic-track duration** shortened by one year, holding fixed **academic content and instruction time** graduation requirements
- Following the reform, each school year has **more hours of instruction** and covers **more academic content**, leading to unambiguous **increase in intensity of schooling process**
- G8 as a **quasi-experiment** because driven by concerns over demographic changes and labor market conditions in Germany, instead of concerns over the schooling process per se

Impact of schooling intensity on (which) learning outcomes

- **End of high school**: shorter duration and higher intensity effects
- **Mid-grade outcomes**: intensity effects only

- For our analysis, mid-grade is **grade nine**, when students are assessed in the German extension of PISA (2000-2012)

Main findings

1. **Higher schooling intensity** has **significant and positive impact** on average scores for the three subjects tested
2. **Heterogeneous effects** across students: girls and students with German born parents, or having more books at home benefit more
3. **Mechanisms**: Effects not explained by changes in observed channels (e.g., out-of-school activities, teacher and classroom quality)
4. Quantile regression results suggest **unobserved heterogeneity** (students' capability to cope with increased intensity) is important

Background: Secondary education in Germany

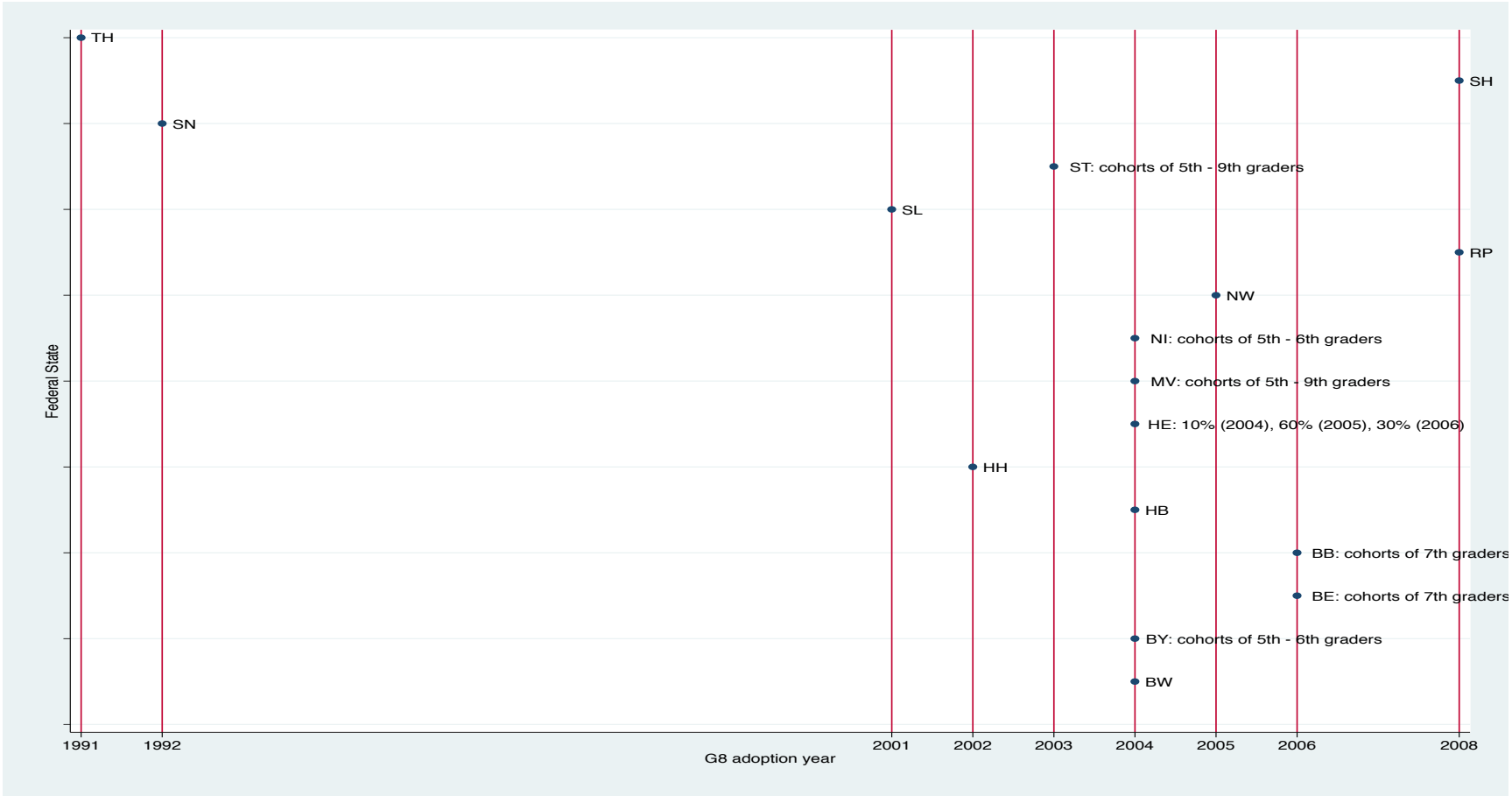
- **Educational policy**: responsibility of the 16 federal states
- **Primary school enrollment**: age 6
- **Primary school length**: four grades
- **Tracking into secondary school**: grade 5
  1. basic-track (*Hauptschule*): grades 5-9
  2. middle-track (*Realschule*): grades 5-10
  3. academic-track (*Gymnasium*): grades 5-13(12): leads to *Abitur*

  - G8 affects only **academic-track curriculum**

Background: The G8 reform

- **Before 2001**, in all but two German states, the academic track lasted nine years: thirteen years of schooling up to graduation
- **Since 2001**, fourteen states started to implement the G8 reform
- While most states began the reform on the **entering student cohort**, a few states extended it to **cohorts already enrolled**

G8 reform implementation timing and first treated cohorts



BW: Baden - Württemberg, **BY**: Bavaria, **BE**: Berlin, **BB**: Brandenburg, **HB**: Bremen, **HH**: Hamburg, **HE**: Hesse, **MV**: Mecklenburg - Vorpommern, **NI**: Lower Saxony, **NW**: North Rhine - Westfalia,**RP**: Rhineland - Palatinate, **SL**: Saarland, **ST**: Saxony - Anhalt, **SN**: Saxony, **SH**: Schleswig - Holstein, **TH**: Thuringia. When not stated otherwise, the first G8 cohorts are fifth graders.

Background: The G8 Reform

1. High school length shortened from 9 (G9) to 8 (G8) grades
2. Overall curriculum kept unaltered
   - same total amount of year-week hours of instruction (265), but now distributed across 8 (rather then 9) grades
   - **more hours of instruction** per week (grade)
   - **more curriculum covered** per week (grade)
   - increase in **learning intensity**

Data: PISA

1. PISA 2000-2012 grade-9 German extension data
   - In each PISA cycle, a range of relevant skills and competencies are assessed in three subjects: Reading, mathematics, and science
   - Using item response theory, PISA maps student achievement in each subject on a standardized scale
   - Our samples include about 34 thousand observations in reading, and about 30 thousand observations in mathematics and science
2. Merged with KulturMinisterKonferenz (KMK) official timetables

Summary statistics

Variable	Mean	SD
PISA scores		
Reading	572.13	55.51
Mathematics	578.39	58.26
Science	587.05	61.10
Student controls		
Female	0.53	0.50
Age (in months)	185.22	5.54
Parental education: Tertiary (ISCED ≥ 5)	0.62	0.49
Parental ISEI	59.25	17.34
Books in house: > 100	0.58	0.49
Only child	0.29	0.45
Foreign born child	0.04	0.20
Foreign born parent	0.13	0.34
Foreign language spoken at home	0.04	0.20
School controls		
School enrollment	793.93	352.15
% of girls enrolled	49.42	15.07
Urban school	0.26	0.44
Private school	0.08	0.26
Student-teacher ratio	14.66	5.93
Student-computer ratio	26.78	62.84
Fraction of certified teachers	0.74	0.40
Fraction of part time teachers	0.35	0.19
Shortage of language arts teachers	0.06	0.24
Shortage of math teachers	0.20	0.40
Shortage of science teachers	0.24	0.43
Shortage of materials for instruction	0.23	0.42
Shortage of lab equipment	0.37	0.48
Shortage of library resources	0.31	0.46
Policy variables		
G8 reform	0.41	0.49
Years of treatment	1.61	2.30
Avg. weekly instruction hours (KMK: grades 5-9)	30.93	1.49

DiD model

- G8 staggered implementation **over time** and **across states** allows use of **difference-in-difference (DiD)** for identification:

$$zscore_{ist} = \beta \cdot G8_{st} + \alpha \cdot \mathbf{X_{ist}} + \delta_s + \gamma_t + \varepsilon_{ist}$$

  - $zscore_{ist}$: standardized PISA score for student i in state s
  - $G8_{st}$: indicator variable for the G8 reform status. Equals one if state s in year t has student cohort treated by G8 reform, zero otherwise
  - X_{ist} : vector of student and school controls
  - δ_s : state fixed effects
  - γ_t : year fixed effects
  - ε_{ist} : residual error term

Treatment definitions

1. **G8 dummy**: Students cohorts coded as treated are typically those subject to the G8 reform upon entering the academic track, except in few states where the reform affected also already enrolled cohorts
2. **Years of treatment**: For some cohorts in some state, length of treatment is shorter than modal treatment duration (5 years)
3. **Year-week hours of instruction**, averaged over grades five through nine by state and cohort (KMK official historical timetables)

Average effects of the G8 reform

	(1)	(2)	(3)
Panel A: Reading ($N = 33,996$)			
G8 reform	0.073***	0.081***	0.078***
	(0.022)	(0.021)	(0.022)
Years of treatment	0.013**	0.015***	0.014**
	(0.005)	(0.005)	(0.006)
Weekly instruction hours	0.031***	0.034***	0.034***
	(0.010)	(0.008)	(0.009)
Panel B: Math ($N = 29,929$)			
G8 reform	0.075*	0.081**	0.067**
	(0.044)	(0.035)	(0.032)
Years of treatment	0.015	0.016*	0.013*
	(0.010)	(0.008)	(0.008)
Weekly instruction hours	0.023*	0.024**	0.022**
	(0.014)	(0.011)	(0.010)
Panel C: Science ($N = 30,202$)			
G8 reform	0.088***	0.089***	0.085***
	(0.026)	(0.020)	(0.018)
Years of treatment	0.017***	0.016***	0.015***
	(0.006)	(0.005)	(0.004)
Weekly instruction hours	0.026**	0.027***	0.025***
	(0.011)	(0.009)	(0.008)
State and year fixed effects	✓	✓	✓
Student controls		✓	✓
School controls			✓

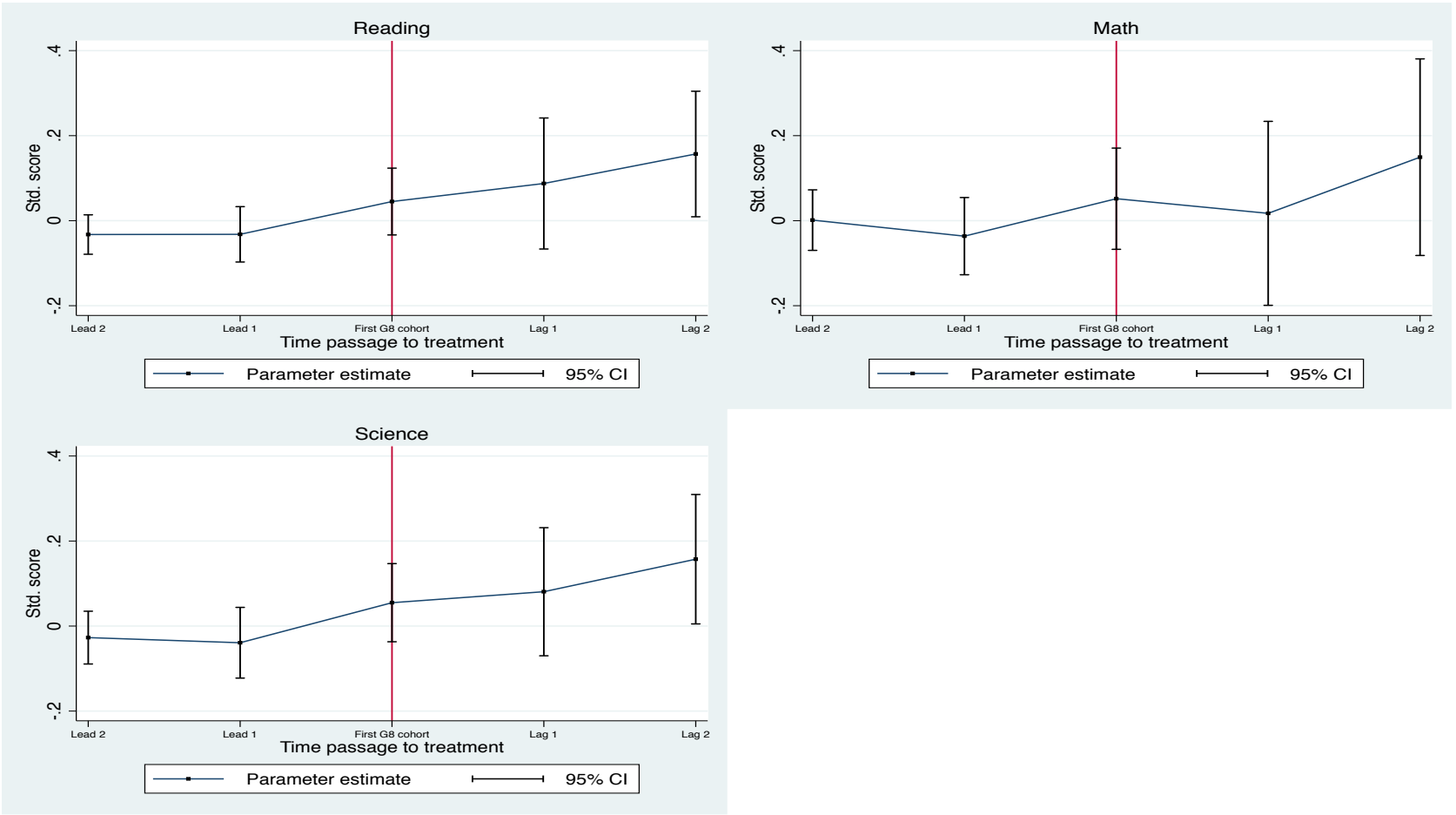
Internal validity of our DiD results

1. Treated and control states should follow **common trends** in the absence of the reform
2. The reform should not induce significant **compositional changes** in the student body
3. **Other contemporaneous reforms** should not have a differential impact on students across treated and control states

Common trends

- **Inter-temporal reform effects** captured by a set of indicators:
  - One for the **first treated cohort**
  - Two **lead variables** (three-year prior and six-year prior)
  - Two **lagged variables** (three-year after and six-or-more-year after)
  - **Omitted category**: nine-or-more-year prior to first treated cohort

Inter-temporal effects of the G8 reform



Source: Computations on PISA 2000-2012 pooled data (baseline specification, final student weights used)

Compositional changes

- No evidence of significant reform impact on observed student and school characteristics
- Consistent with the stability of main results across specifications 1-3

Contemporaneous education policy changes

- Add a dummy indicating the state-specific cohorts affected by the introduction of **Central Exit Examinations** (6)
- Control for share of all-day students in a state, to account for federal investment program promoting introduction of **all-day schooling** (7)

Sensitivity analysis

	Main spec. (1)	DD lead (2)	Placebos lower-tracks (3)	DDD model (4)	State trends (5)	Switch to CEE (6)	All day schooling (7)	Double cohorts (8)
Panel A: Reading								
G8 reform	0.078***	-0.010	-0.017	0.115**	0.075**	0.069**	0.077***	0.081***
	(0.022)	(0.026)	(0.045)	(0.058)	(0.029)	(0.031)	(0.022)	(0.023)
Observations	33,996	33,996	57,748	72,053	33,996	33,996	33,996	33,996
Panel B: Math								
G8 reform	0.067**	-0.036	-0.022	0.092*	0.100***	0.061*	0.065**	0.065**
	(0.032)	(0.041)	(0.058)	(0.054)	(0.034)	(0.050)	(0.032)	(0.033)
Observations	29,929	29,929	50,542	63,289	29,929	29,929	29,929	29,929
Panel C: Science								
G8 reform	0.085***	-0.020	-0.001	0.094**	0.097***	0.084***	0.084***	0.089***
	(0.018)	(0.038)	(0.058)	(0.039)	(0.032)	(0.033)	(0.018)	(0.017)
Observations	30,202	30,202	50,988	63,886	30,202	30,202	30,202	30,202

Further robustness checks

1. Estimate **placebo regressions** where the reform dummy equals one for the cohort immediately before the first cohort actually treated (2)
2. Consider the reform impact on basic- and middle-track students (3)
3. Middle-track students as additional control group in **DDD** (4)
4. Allow treated and control states to follow different **linear trends** (5)
5. Include a dummy for **double graduating cohorts** (8)

Sensitivity analysis

	Main spec. (1)	DD lead (2)	Placebos lower-tracks (3)	DDD model (4)	State trends (5)	Switch to CEE (6)	All day schooling (7)	Double cohorts (8)
Panel A: Reading								
G8 reform	0.078***	-0.010	-0.017	0.115**	0.075**	0.069**	0.077***	0.081***
	(0.022)	(0.026)	(0.045)	(0.058)	(0.029)	(0.031)	(0.022)	(0.023)
Observations	33,996	33,996	57,748	72,053	33,996	33,996	33,996	33,996
Panel B: Math								
G8 reform	0.067**	-0.036	-0.022	0.092*	0.100***	0.061*	0.065**	0.065**
	(0.032)	(0.041)	(0.058)	(0.054)	(0.034)	(0.050)	(0.032)	(0.033)
Observations	29,929	29,929	50,542	63,289	29,929	29,929	29,929	29,929
Panel C: Science								
G8 reform	0.085***	-0.020	-0.001	0.094**	0.097***	0.084***	0.084***	0.089***
	(0.018)	(0.038)	(0.058)	(0.039)	(0.032)	(0.033)	(0.018)	(0.017)
Observations	30,202	30,202	50,988	63,886	30,202	30,202	30,202	30,202

Heterogeneous reform effects (observed heterogeneity)

- Allowing different reform effects for different subgroups of students:

$$zscore_{ist} = \sum_{g=1}^{N} \beta_g \cdot G8_{st} \cdot I(\text{student } i \in \text{ subgroup } g) + \alpha \mathbf{X_{ist}} + \delta_s + \gamma_t + \varepsilon_{ist}$$

  - $I(\cdot) = 1$ if student i belongs in one of N subgroups indexed by g
  - β_g : reform effect for a subgroup defined by a student characteristic

Heterogeneous reform effects based on observed heterogeneity

	Reading (1)	Mathematics (2)	Science (3)
Panel A: Reform effects by gender			
G8 × Boy	0.003	0.081*	0.065***
	(0.024)	(0.043)	(0.020)
G8 × Girl	0.143***	0.056*	0.101***
	(0.025)	(0.030)	(0.020)
P − value of t-test of difference	**0.00**	0.50	**0.05**
Observations	33,922	29,885	30,128
Panel B: Reform effects by age			
G8 × Age ≤ 1st tercile	0.059**	0.044	0.062**
	(0.027)	(0.034)	(0.026)
G8 × Age > 1st tercile	0.094***	0.087***	0.104***
	(0.027)	(0.033)	(0.021)
P − value of t-test of difference	0.22	**0.02**	0.16
Observations	33,996	29,929	30,202
Panel C: Reform effects by parental immigration status			
G8 × Foreign born parents	0.019	0.007	-0.045
	(0.063)	(0.055)	(0.052)
G8 × German born parents	0.075***	0.060**	0.087***
	(0.020)	(0.030)	(0.018)
P − value of t-test of difference	0.29	0.19	**0.00**
Observations	33,126	29,058	29,331
Panel D: Reform effects by parental education			
G8 × < Tertiary	0.099**	0.071	0.095***
	(0.038)	(0.044)	(0.033)
G8 × Tertiary	0.061***	0.052*	0.071***
	(0.023)	(0.027)	(0.023)
P − value of t-test of difference	0.36	0.60	0.54
Observations	32,861	28,793	29,066
Panel E: Reform effects by parental ISEI			
G8 × ISEI ≤ 1st tercile	0.076**	0.040	0.073***
	(0.030)	(0.037)	(0.024)
G8 × ISEI > 1st tercile	0.079***	0.083**	0.091***
	(0.022)	(0.032)	(0.019)
P − value of t-test of difference	0.88	**0.10**	0.40
Observations	33,680	29,612	29,885
Panel F: Reform effects by book at home			
G8 × Books at home: ≤ 100	0.050*	0.028	0.020
	(0.028)	(0.042)	(0.032)
G8 × Books at home: > 100	0.088***	0.076**	0.105***
	(0.022)	(0.031)	(0.019)
P − value of t-test of difference	0.09	0.14	**0.01**
Observations	32,774	28,765	29,032

Mechanisms

- By adding more instruction time in school, the reform could
  1. Reduce time available for **out-of-school activities**, especially those that are academically productive such as homework, extracurricular programs, and/or remedial work for struggling students
  2. Affect **teacher quality** if composition of high school teachers changes
  3. Affect **classroom quality** if the higher intensity increases the stress level for teachers and/or students, leading to behavioral changes in the interactive classroom environment
- We find no significant changes after the reform for these channels

Effects on out-of-school study time/attendance and teachers/classroom quality

	Indexes of:			
	Out-of-school study time (1)	**Out-of-school class attendance (2)**	**Teacher quality (3)**	**Classroom quality (4)**
G8 reform	-0.053	-0.052	0.056	0.201
	(0.157)	(0.149)	(0.044)	(0.138)
Observations	7,973	9,162	29,081	21,574

Heterogeneous reform effects (unobserved heterogeneity)

- Allowing different reform effects for students at different quantiles of the conditional test score distribution:

$$h_{\tau,st} = \beta_\tau G8_{st} + \alpha_\tau \mathbf{X_{ist}} + \delta_{\tau,s} + \gamma_{\tau,t}$$

  - $h_{\tau,st}$: test score at the τ-th quantile of the distribution in state s and year t
  - β_τ : reform effects on students at different quantiles of conditional test score distribution according to their unobserved heterogeneity

Heterogeneous effects of the G8 reform: unobserved heterogeneity

	Quantiles								
	0.10	**0.20**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
Panel A: Reading ($N = 33,996$)									
G8	0.038	0.054*	0.065**	0.075**	0.092***	0.098***	0.097***	0.102***	0.104***
	(0.036)	(0.032)	(0.029)	(0.031)	(0.025)	(0.023)	(0.025)	(0.032)	(0.040)
Panel B: Math ($N = 29,929$)									
G8	0.037	0.025	0.033	0.062***	0.072***	0.091***	0.093***	0.099***	0.084**
	(0.038)	(0.037)	(0.038)	(0.024)	(0.026)	(0.033)	(0.028)	(0.029)	(0.042)
Panel C: Science ($N = 30,202$)									
G8	0.059	0.071**	0.063**	0.077***	0.083***	0.096***	0.098***	0.102***	0.101**
	(0.047)	(0.036)	(0.028)	(0.025)	(0.028)	(0.024)	(0.024)	(0.029)	(0.048)

Conclusions

- We estimate the impact of an increase in schooling intensity
- Using PISA 2000-2012 data, we find that average test scores improve significantly as a consequence of the increase
- Reform effects are small, and **heterogeneous** across students
  - along **observed dimensions**: girls, students with German born parents, and having more books at home benefit more
  - along **the unobserved dimension**: high-performing students benefit more than low-performing students
- Our results have **implications** beyond the German context:
  - for countries that are considering similar reforms (e.g., Italy)
  - at a more disaggregate level, when schooling intensity may change as a result of personal choices (e.g., part-time versus full-time college)