No country for young kids?

The effects of school starting age throughout childhood and beyond

Gonçalo Lima

Nova School of Business and Economics

goncalo.lima@novasbe.pt

Co-authors:

Miguel Madeira Ruivo	Ana Balcão Reis	Luís Catela Nunes	Maria do Carmo Seabra
Nova SBE	Nova SBE	Nova SBE	Nova SBE

Oxford Education Research Symposium - July 30, 2019

Introduction

Motivation

Every year many children enter kindergarten or school for the first time:

Introduction

Motivation

Every year many children enter kindergarten or school for the first time:

- Enter at different stages of development

Introduction

Motivation

Every year many children enter kindergarten or school for the first time:

- Enter at different stages of development
- Differences in age tend to have strong impacts in early childhood

Introduction

Motivation

Every year many children enter kindergarten or school for the first time:

- Enter at different stages of development
- Differences in age tend to have strong impacts in early childhood
- Early childhood differences $\rightarrow$ long-term impacts
  $\longrightarrow$ e.g. Chetty et al. (2011); Heckman (2011)

Introduction

Motivation

Relevance for policy and parents:

- Changing age cutoffs
  - $\longrightarrow$ Bedard and Dhuey (2006); Deming and Dynarski (2008); Elder and Lubotsky (2009)

Introduction

Motivation

Relevance for policy and parents:

- Changing age cutoffs
  → Bedard and Dhuey (2006); Deming and Dynarski (2008); Elder and Lubotsky (2009)
- School entrance postponement
  → Deming and Dynarski (2008); Cook and Kang (2018)

Introduction

Motivation

Relevance for policy and parents:

- Changing age cutoffs
  - → Bedard and Dhuey (2006); Deming and Dynarski (2008); Elder and Lubotsky (2009)
- School entrance postponement
  - → Deming and Dynarski (2008); Cook and Kang (2018)

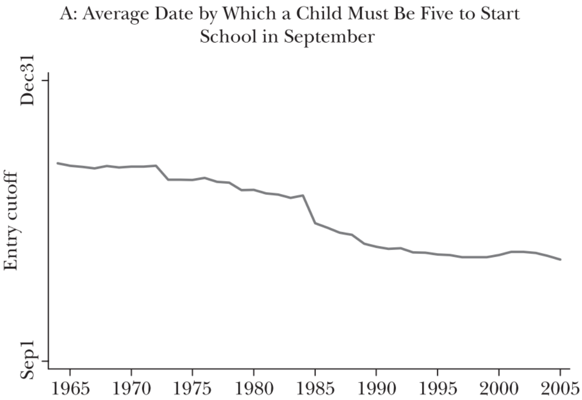


Figure reproduced from Deming and Dynarski (2008)

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects
- Add evidence to literature from previously unstudied context ▸ Literature

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects
- Add evidence to literature from previously unstudied context ▸ Literature
- Leverage exact birth date variations close to cutoff

  $\longrightarrow$ McEwan and Shapiro (2008); Evans et al. (2010); Dobkin and Ferreira (2010); Peña (2017); Attar and Cohen-Zada (2018)

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects
- Add evidence to literature from previously unstudied context ▸ Literature
- Leverage exact birth date variations close to cutoff
  - $\longrightarrow$ McEwan and Shapiro (2008); Evans et al. (2010); Dobkin and Ferreira (2010); Peña (2017); Attar and Cohen-Zada (2018)
- Analyze long-term education outcomes

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects
- Add evidence to literature from previously unstudied context ▸ Literature
- Leverage exact birth date variations close to cutoff
  → McEwan and Shapiro (2008); Evans et al. (2010); Dobkin and Ferreira (2010); Peña (2017); Attar and Cohen-Zada (2018)
- Analyze long-term education outcomes
- Go beyond intent-to-treat (ITT) causal effects

Introduction

Contributions

- Provide evidence with high-stakes policy outcomes
- Use quasi-experiment to identify causal effects
- Add evidence to literature from previously unstudied context ▸ Literature
- Leverage exact birth date variations close to cutoff
  - $\longrightarrow$ McEwan and Shapiro (2008); Evans et al. (2010); Dobkin and Ferreira (2010); Peña (2017); Attar and Cohen-Zada (2018)
- Analyze long-term education outcomes
- Go beyond intent-to-treat (ITT) causal effects

Introduction

Main findings

Being 1-year older when starting school leads to:

Introduction

Main findings

Being 1-year older when starting school leads to:

- $\uparrow$ cognitive capacity (Math and Language)

Introduction

Main findings

Being 1-year older when starting school leads to:

- $\uparrow$ cognitive capacity (Math and Language)
- $\downarrow$ probability of repeating

Introduction

Main findings

Being 1-year older when starting school leads to:

- $\uparrow$ cognitive capacity (Math and Language)
- $\downarrow$ probability of repeating

Persistence of cognitive gains is limited:

- Cognitive effects fade quickly

Introduction

Main findings

Being 1-year older when starting school leads to:

- $\uparrow$ cognitive capacity (Math and Language)
- $\downarrow$ probability of repeating

Persistence of cognitive gains is limited:

- Cognitive effects fade quickly
- Limited heterogeneity

Introduction

Main findings

Being 1-year older when starting school leads to:

- $\uparrow$ cognitive capacity (Math and Language)
- $\downarrow$ probability of repeating

Persistence of cognitive gains is limited:

- Cognitive effects fade quickly
- Limited heterogeneity
- Driven by differences in cognitive maturity

Effects persist through institutional features:

- $\downarrow$ probability of dropping out

Introduction

Preliminary findings

Effects persist through institutional features:

- $\downarrow$ probability of dropping out
- $\uparrow$ probability of graduating

Effects persist through institutional features:

- $\downarrow$ probability of dropping out
- $\uparrow$ probability of graduating
- $\uparrow$ probability of enrolling in academic track

Introduction

Preliminary findings

Effects persist through institutional features:

- $\downarrow$ probability of dropping out
- $\uparrow$ probability of graduating
- $\uparrow$ probability of enrolling in academic track
- $\uparrow$ probability of enrolling in scientific curricula

Introduction

Preliminary findings

Effects persist through institutional features:

- $\downarrow$ probability of dropping out
- $\uparrow$ probability of graduating
- $\uparrow$ probability of enrolling in academic track
- $\uparrow$ probability of enrolling in scientific curricula
- $\uparrow$ application scores to public HE

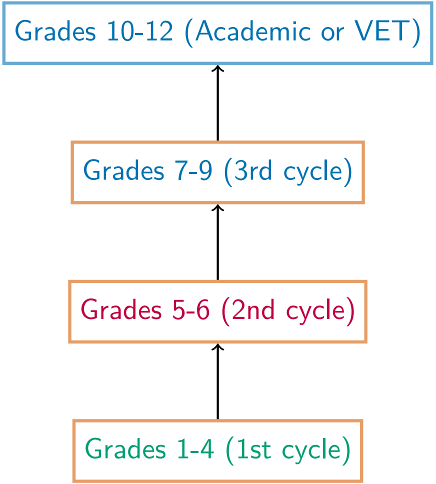
Introduction

Preliminary findings

Effects persist through institutional features:

- $\downarrow$ probability of dropping out
- $\uparrow$ probability of graduating
- $\uparrow$ probability of enrolling in academic track
- $\uparrow$ probability of enrolling in scientific curricula
- $\uparrow$ application scores to public HE
- $\uparrow$ probability of enrolling in more selective public HE courses

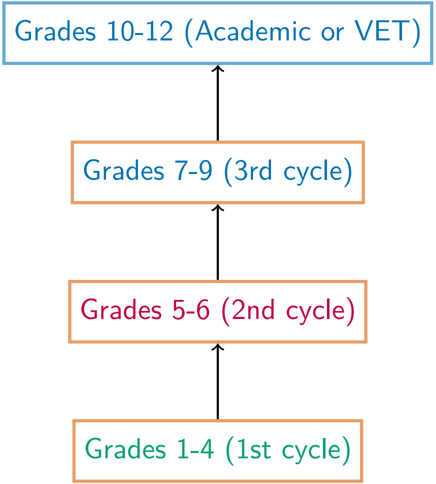
Institutional background



Relevant institutional features:

- National exams at the end of each cycle

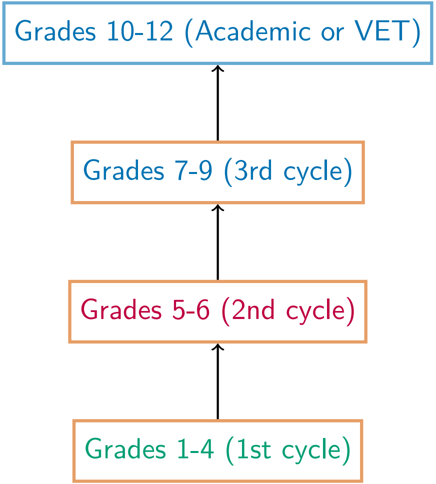
Institutional background



Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9

Institutional background

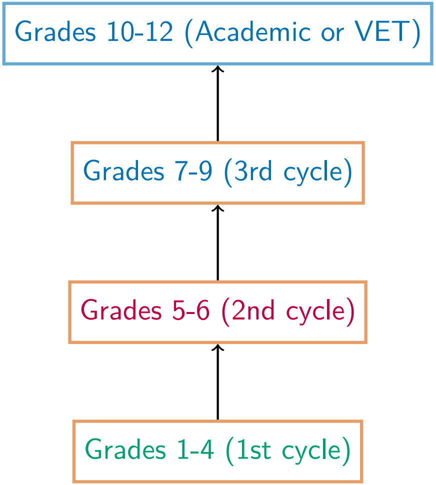


Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9

  $\rightarrow$ Multiple VET courses + 4 academic streams

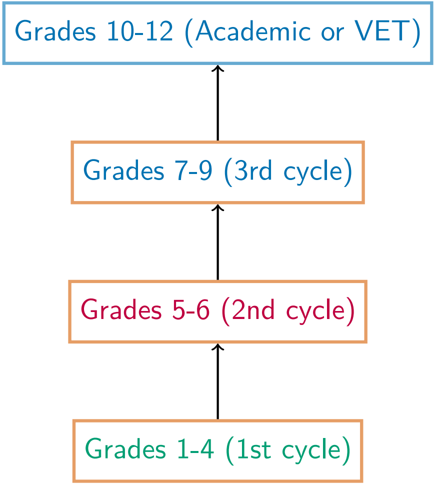
Institutional background



Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9
  $\rightarrow$ Multiple VET courses $+$ 4 academic streams
- Compulsory schooling until 18-years of age

Institutional background



Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9
  - $\rightarrow$ Multiple VET courses + 4 academic streams
- Compulsory schooling until 18-years of age
- High grade retention rates

Institutional background

Grades 10-12 (Academic or VET)

↑

Grades 7-9 (3rd cycle)

↑

Grades 5-6 (2nd cycle)

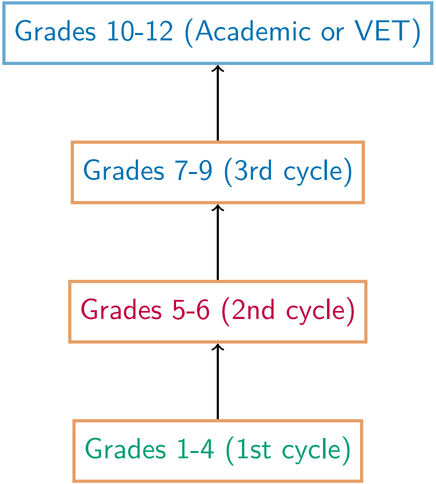
↑

Grades 1-4 (1st cycle)

Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9
  $\rightarrow$ Multiple VET courses + 4 academic streams
- Compulsory schooling until 18-years of age
- High grade retention rates
- HE access based on exit exams and GPA

Institutional background



Relevant institutional features:

- National exams at the end of each cycle
- Multi-pronged tracking at end of Grade 9
  $\rightarrow$ Multiple VET courses + 4 academic streams
- Compulsory schooling until 18-years of age
- High grade retention rates
- HE access based on exit exams and GPA
- HE divided in academic and polytechnic offer

Institutional background

School entry laws:

Institutional background

School entry laws:

- 6-years old by 15 September

Institutional background

School entry laws:

- 6-years old by 15 September
- Born 16Sept - 31Dec: can still enroll

Institutional background

School entry laws:

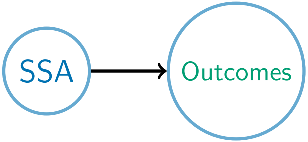
- 6-years old by 15 September
- Born 16Sept - 31Dec: can still enroll
- Binding cutoff at 1 January

Empirical strategy

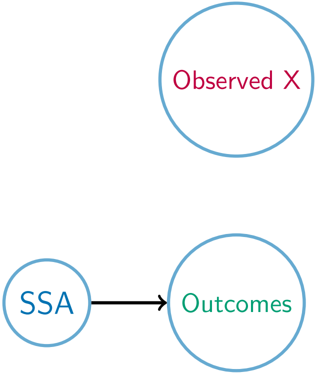
Quasi-experiment:

- Random variation induced by birth dates
- Distance in days to cutoff predicts SSA
- Plausibly causal effects
- Local polynomial estimates ‣ Show me the math

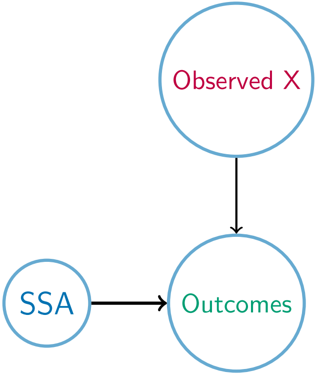
Empirical strategy



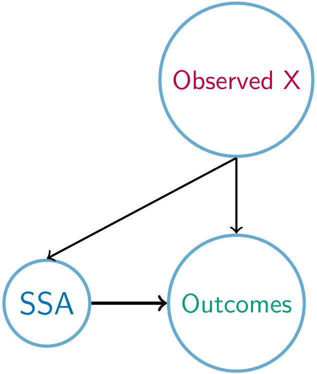
Empirical strategy



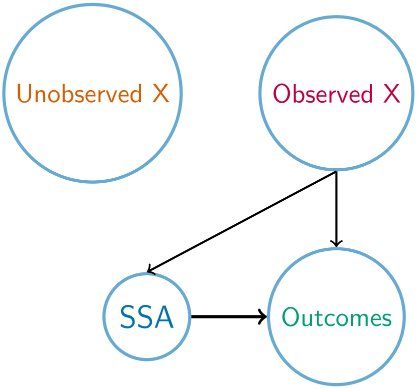
Empirical strategy



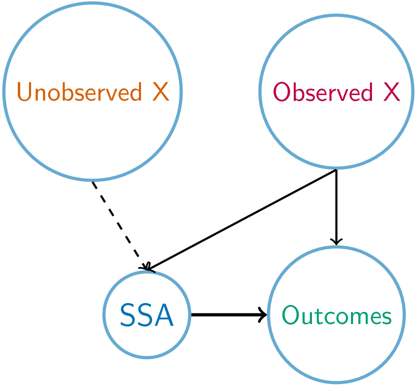
Empirical strategy



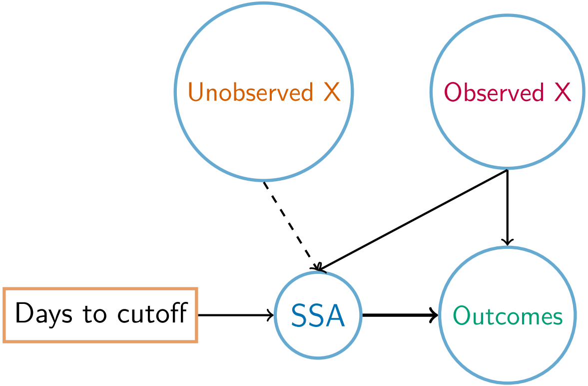
Empirical strategy



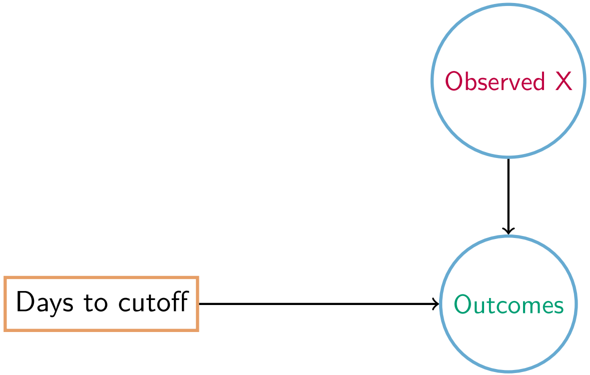
Empirical strategy



Empirical strategy



Empirical strategy



Data

Description

Administrative data from Portugal:

Data

Description

Administrative data from Portugal:

- Data on every student and teacher (Grades 1-12)

Data

Description

Administrative data from Portugal:

- Data on every student and teacher (Grades 1-12)
- Data on national exam scores + internal scores

Data
Description

Administrative data from Portugal:

- Data on every student and teacher (Grades 1-12)
- Data on national exam scores + internal scores
- Analytical datasets:

Data
Description

Administrative data from Portugal:

- Data on every student and teacher (Grades 1-12)
- Data on national exam scores + internal scores
- Analytical datasets:
  - $\rightarrow$ Dataset 1, Grades 1-9: born 1998-2008; sat exams 2012-2017; **660k obs.**
  - $\rightarrow$ Dataset 2, Grades 9-12: born 1988-2000; sat exams 2007-2014; **635k obs.**

▸ What about attrition?

Data

Main variables

Outcomes:

- Achievement (Grades 4, 6, 9, 11, 12)
- Grade retention (until Grades 4, 6, 9-12)
- Dropout and graduation (Grade 9)
- Track choice (Grade 10)
- Academic course choice in HS (Grade 10)
- HE application outcomes

Data

Main variables

Outcomes

SSA

Data

Main variables

Outcomes

SSA

Birth date

Data
Main variables

Outcomes

SSA

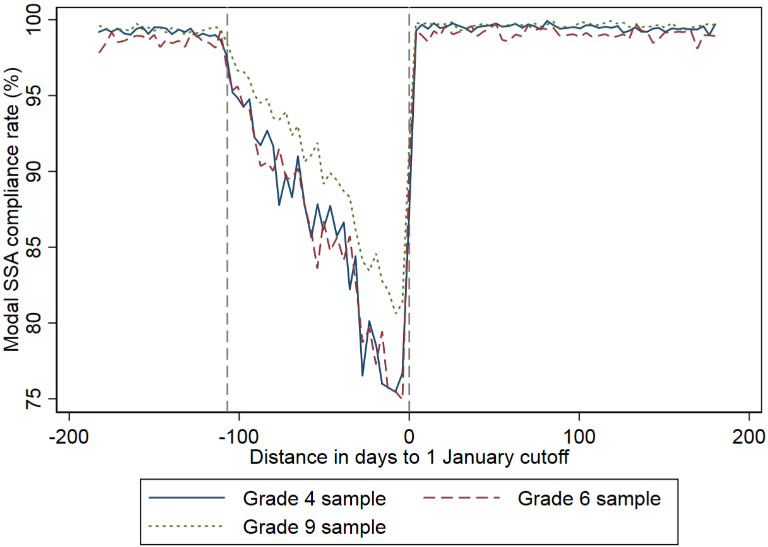
Birth date

Student char.: ‣ Descriptive statistics

- Female
- First generation immigrant
- Computer at home
- School social support
- If dad unemployed
- Household level of education

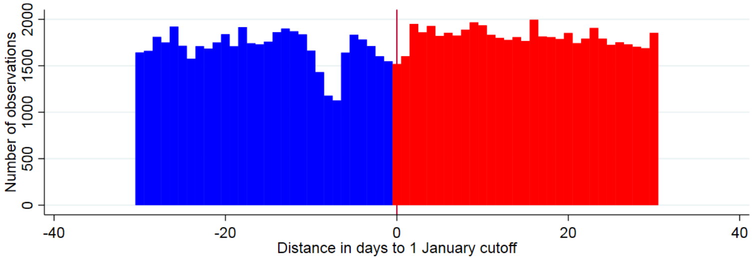
Data

Compliance rates



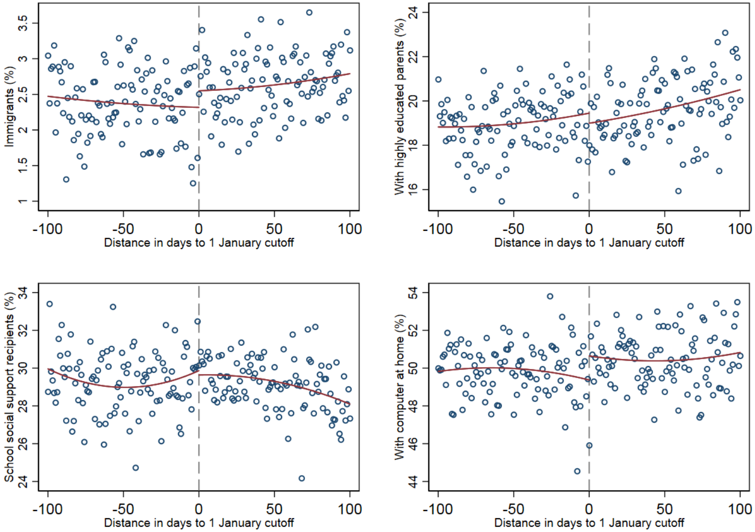
Results

No evidence of birth date manipulation



Results

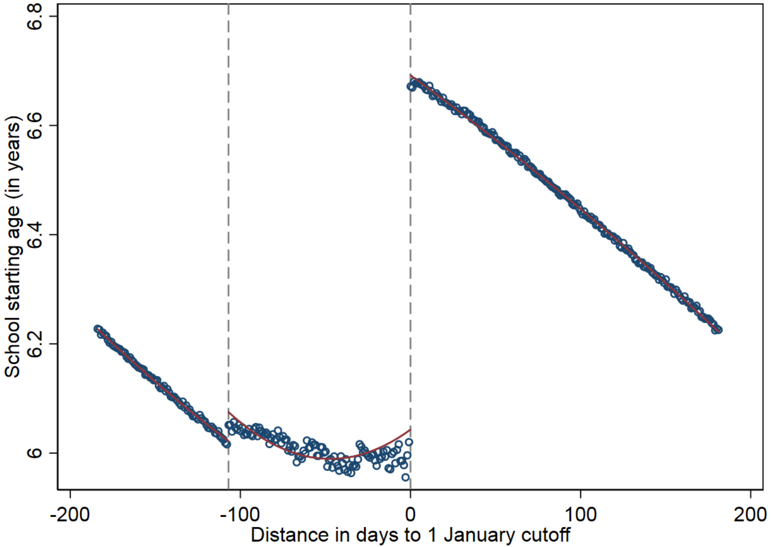
Continuity of covariates at the cutoff ‣ Show me



Results

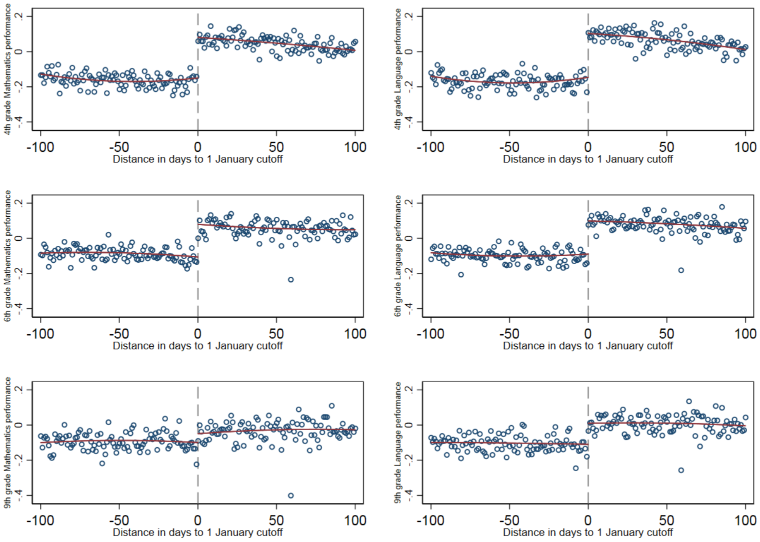
Discontinuity in school starting age

▸ Estimates by grade



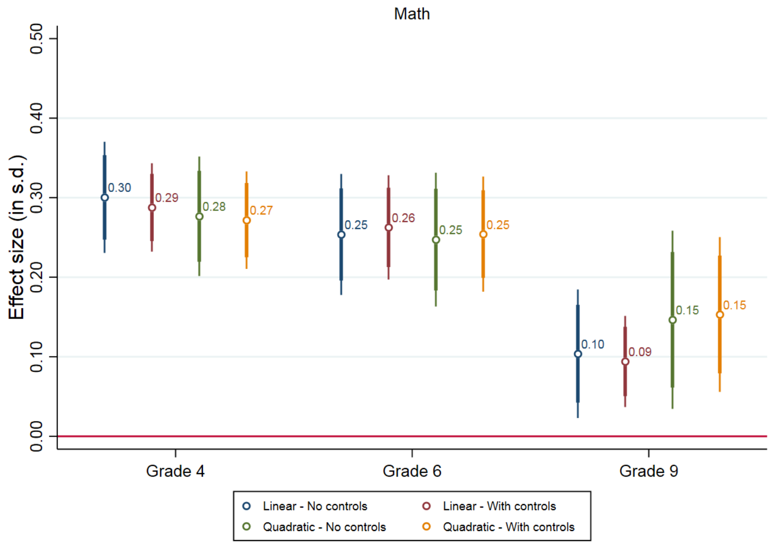
Results

Discontinuities in student achievement



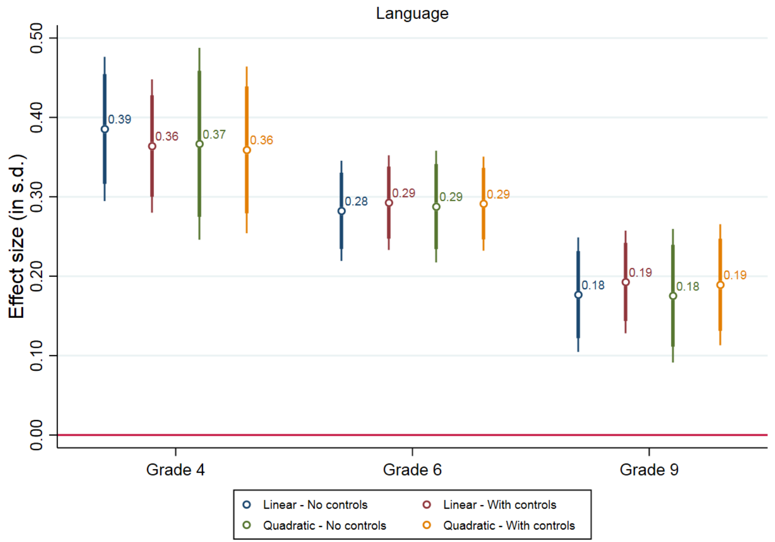
Results

Basic education outcomes → LATE Math ▸ Power



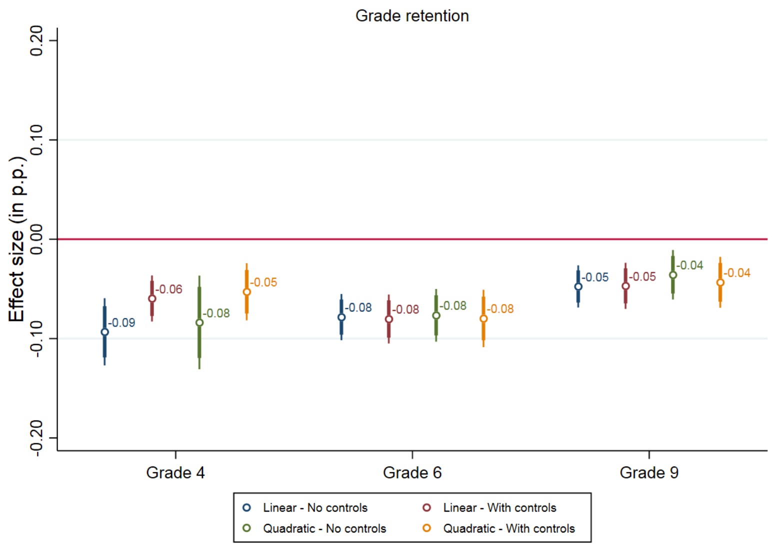
Results

Basic education outcomes → LATE Language



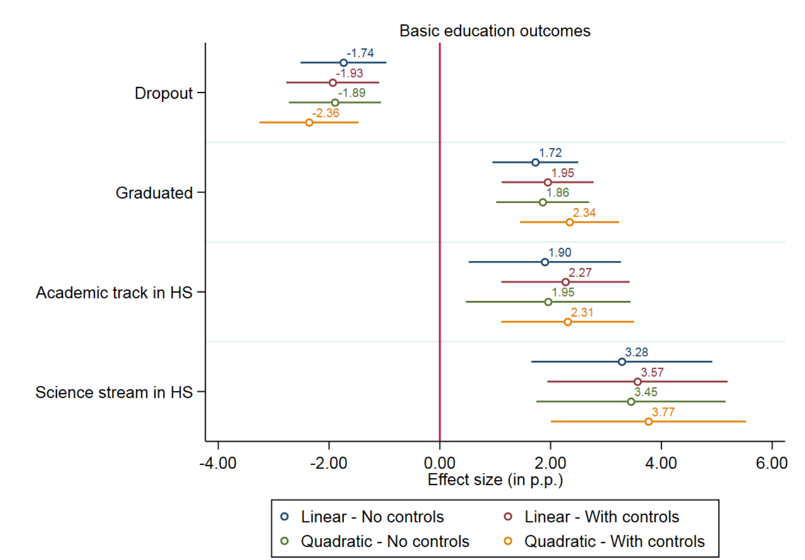
Results

Basic education outcomes → LATE Grade retention



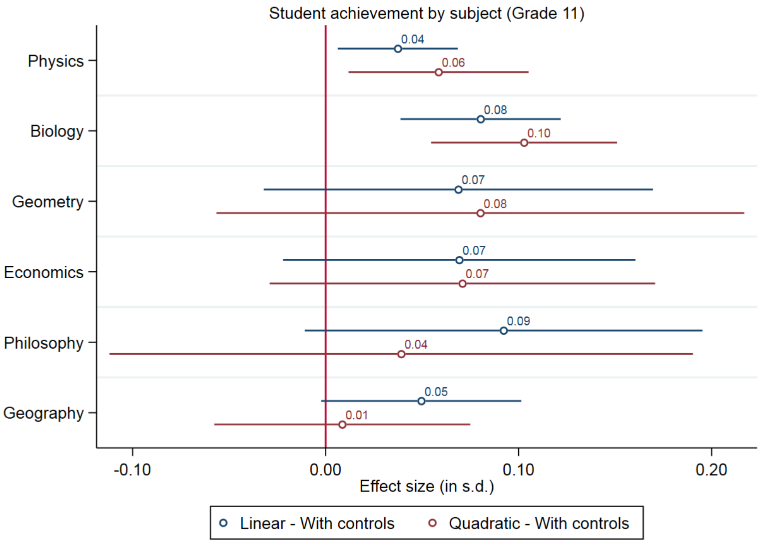
Results

Grade 9 attainment outcomes → ITT



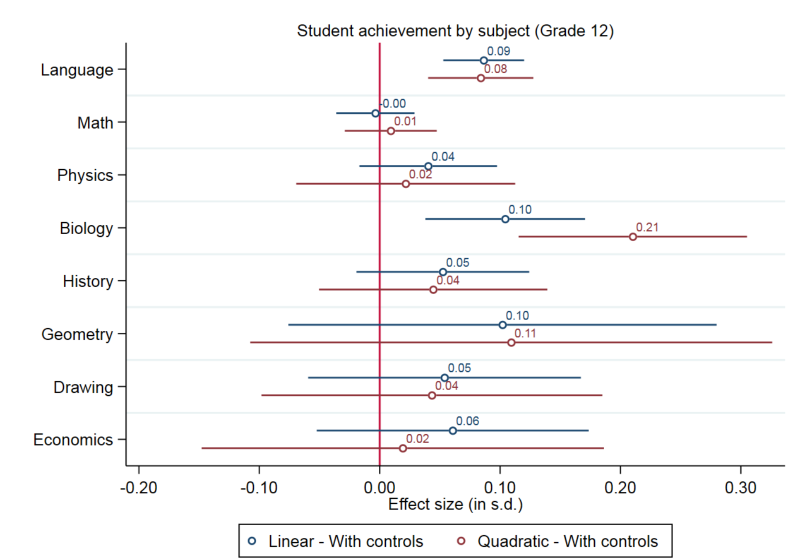
Results

High-school outcomes → ITT Grade 11 Achievement



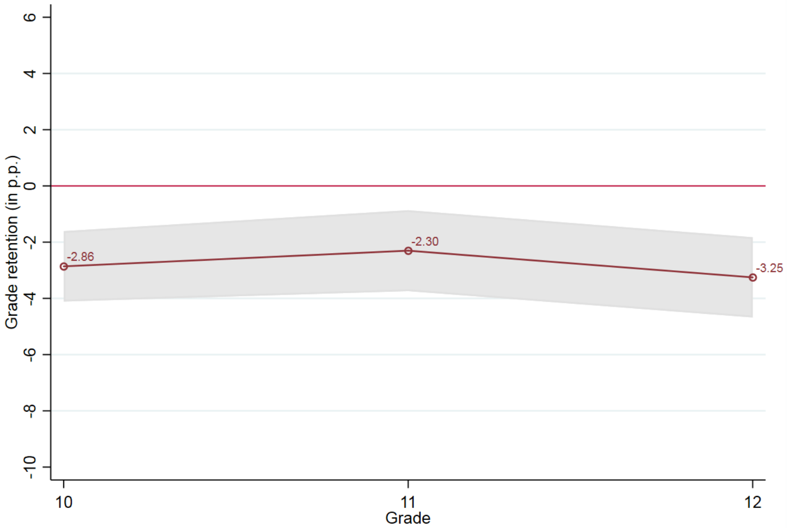
Results

High-school outcomes → ITT Grade 12 Achievement



Results

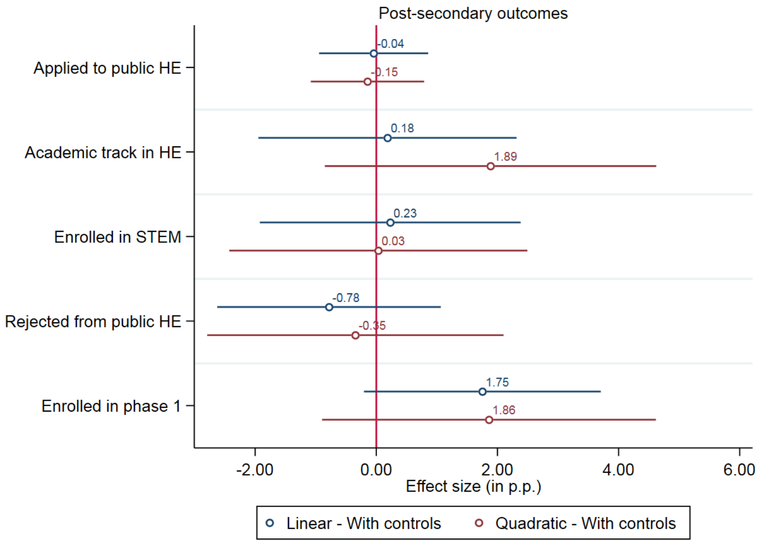
High school outcomes → ITT Grade retention in high school



Shadowed area: 95% confidence intervals

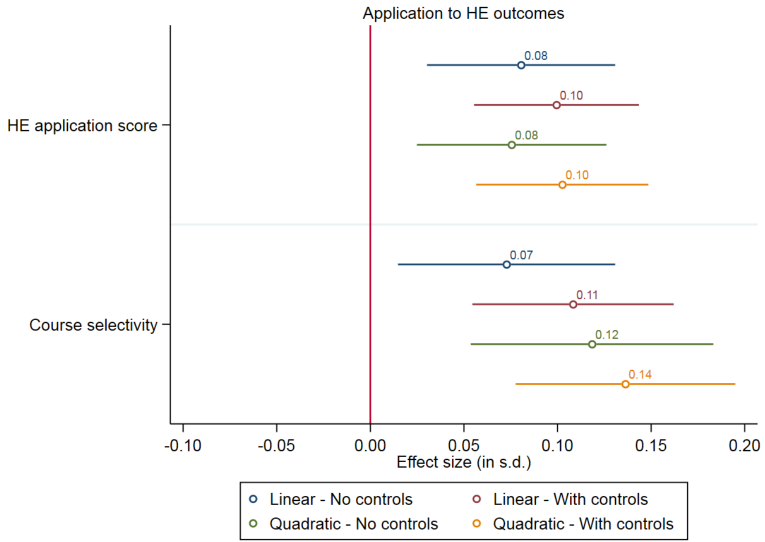
Results

Post-secondary outcomes → ITT application to HE outcomes



Results

Post-secondary outcomes → ITT application to HE outcomes



Results

Robustness checks

Point estimates are stable across:

- Regressions only with non-repeaters ‣ Show me

Results

Robustness checks

Point estimates are stable across:

- Regressions only with non-repeaters ‣ Show me
- Regressions with birth day of the week FEs ‣ Show me

Results

Robustness checks

Point estimates are stable across:

- Regressions only with non-repeaters ‣ Show me
- Regressions with birth day of the week FEs ‣ Show me
- Regressions with alternative bandwidths ‣ Show me

Results

Robustness checks

Point estimates are stable across:

- Regressions only with non-repeaters ‣ Show me
- Regressions with birth day of the week FEs ‣ Show me
- Regressions with alternative bandwidths ‣ Show me
- Placebo specifications and permutation-based p-values ‣ Show me
- ‣ Discussion of mechanisms

Conclusion

Being 1-year older has many individual short-term benefits:

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12
- $\uparrow$ likelihood of choosing academic track and science-related curricula

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12
- $\uparrow$ likelihood of choosing academic track and science-related curricula
- $\uparrow$ HE application score

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12
- $\uparrow$ likelihood of choosing academic track and science-related curricula
- $\uparrow$ HE application score
- $\uparrow$ likelihood of enrolling in more selective HE courses

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12
- $\uparrow$ likelihood of choosing academic track and science-related curricula
- $\uparrow$ HE application score
- $\uparrow$ likelihood of enrolling in more selective HE courses
- No evidence of differences in rejection rates and enrollment in STEM

Conclusion

Being 1-year older has many individual short-term benefits:

- $\uparrow$ achievement: LATE always above $.15\sigma$ until Grade 9
  - $\rightarrow$ Similar to Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Attar and Cohen-Zada (2018)
- $\downarrow$ likelihood of grade retention: about 6pp until Grade 9, 3pp until Grade 12
- $\uparrow$ likelihood of choosing academic track and science-related curricula
- $\uparrow$ HE application score
- $\uparrow$ likelihood of enrolling in more selective HE courses
- No evidence of differences in rejection rates and enrollment in STEM
- Cognitive effects fade quickly, institutional features ensure persistence

Conclusion

Parental strategy may be at odds with policy prescription:

Conclusion

Parental strategy may be at odds with policy prescription:

- Delayed enrollment brings individual benefits
- Unfair to parents constrained in their choice

Conclusion

Parental strategy may be at odds with policy prescription:

- Delayed enrollment brings individual benefits
- Unfair to parents constrained in their choice
- Individual costs: additional pre-school costs shorter work careers

Conclusion

Parental strategy may be at odds with policy prescription:

- Delayed enrollment brings individual benefits
- Unfair to parents constrained in their choice
- Individual costs: additional pre-school costs shorter work careers
- Social costs: distributional effects and no gain in earlier cutoffs

Thank you for your attention!

References I

Attar, I. and D. Cohen-Zada (2018). The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth. *Journal of Economic Behavior and Organization 153*(10568), 38–57.

Bedard, K. and E. Dhuey (2006, nov). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics 121*(4), 1437–1472.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011, may). Too Young to Leave the Nest? The Effects of School Starting Age. *Review of Economics and Statistics 93*(2), 455–467.

Cascio, E. U. and D. W. Schanzenbach (2016, jul). First in the Class? Age and the Education Production Function. *Education Finance and Policy 11*(3), 225–250.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Cook, P. and S. Kang (2018). The School-Entry-Age Rule Affects Redshirting Patterns and Resulting Disparities in Achievement.

References II

Cook, P. J. and S. Kang (2016). Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation. *American Economic Journal: Applied Economics 8*(1), 33–57.

Crawford, C., L. Dearden, and C. Meghir (2007). When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England.

Deming, D. and S. Dynarski (2008). The Lengthening of Childhood. *Journal of Economic Perspectives 22*(3), 71–92.

Dhuey, E., D. Figlio, K. Karbownik, and J. Roth (2017). Age and Cognitive Development.

Dhuey, E. and S. Lipscomb (2008, apr). What makes a leader? Relative age and high school leadership. *Economics of Education Review 27*(2), 173–183.

Dhuey, E. and S. Lipscomb (2010, oct). Disabled or young? Relative age and special education diagnoses in schools. *Economics of Education Review 29*(5), 857–872.

Dobkin, C. and F. Ferreira (2010, feb). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review 29*(1), 40–54.

Du, Q., H. Gao, and M. D. Levi (2012, dec). The relative-age effect and career success: Evidence from corporate CEOs. *Economics Letters 117*(3), 660–662.

References III

Elder, T. E. and D. H. Lubotsky (2009). Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers. *The Journal of Human Resources 44*(3), 641–683.

Evans, W. N., M. S. Morrill, and S. T. Parente (2010). Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children. *Journal of Health Economics 29*(5), 657–673.

Fredriksson, P. and B. Öckert (2014, sep). Life-cycle effects of age at school start. *Economic Journal 124*(579), 977–1004.

Heckman, J. J. (2011). The Economics of Inequality. The Value of Early Childhood Education. *American Educator 35*(1), 31–36.

Landersø, R., H. S. Nielsen, and M. Simonsen (2017, jun). School Starting Age and the Crime-age Profile. *Economic Journal 127*(602), 1096–1118.

McEwan, P. J. and J. S. Shapiro (2008). The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates The Benefits of Delayed Primary School Enrollment Discontinuity Estimates Using Exact Birth Dates. *The Journal of Human Resources 43*(1), 1–29.

Mühlenweg, A., D. Blomeyer, H. Stichnoth, and M. Laucht (2012, jun). Effects of age at school entry (ASE) on the development of non-cognitive skills: Evidence from psychometric data. *Economics of Education Review 31*(3), 68–76.

References IV

Mühlenweg, A. M. and P. A. Puhani (2010). The Evolution of the School-Entry Age Effect in a School Tracking System. *Journal of Human Resources 45*(2), 407–438.

Peña, P. A. (2017, feb). Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood. *Economics of Education Review 56*, 152–176.

Puhani, P. A. and A. M. Weber (2007, may). Does the early bird catch the worm? *Empirical Economics 32*(2-3), 359–386.

Schneeweis, N. and M. Zweimüller (2014, apr). Early Tracking and the Misfortune of Being Young. *The Scandinavian Journal of Economics 116*(2), 394–428.

Appendix

Related literature

Older children have many academic benefits:

Appendix

Related literature

Older children have many academic benefits:

- Score higher, repeat less, stay longer in school
  - $\longrightarrow$ e.g. Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Cascio and Schanzenbach (2016); Attar and Cohen-Zada (2018)

Appendix
Related literature

Older children have many academic benefits:

- Score higher, repeat less, stay longer in school
  - $\longrightarrow$ e.g. Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Cascio and Schanzenbach (2016); Attar and Cohen-Zada (2018)
- More likely tracked into academic curricula
  - $\longrightarrow$ Puhani and Weber (2007); Schneeweis and Zweimüller (2014); Attar and Cohen-Zada (2018)

Appendix

Related literature

Older children have many academic benefits:

- Score higher, repeat less, stay longer in school
  → e.g. Bedard and Dhuey (2006); Puhani and Weber (2007); McEwan and Shapiro (2008); Cascio and Schanzenbach (2016); Attar and Cohen-Zada (2018)
- More likely tracked into academic curricula
  → Puhani and Weber (2007); Schneeweis and Zweimüller (2014); Attar and Cohen-Zada (2018)
- Improve outcomes of their younger peers
  → Cascio and Schanzenbach (2016)

Appendix

Related literature

But also non-cognitive benefits:

Appendix

Related literature

But also non-cognitive benefits:

- Less likely to be classified with ADHD
  - $\longrightarrow$ Dhuey and Lipscomb (2010); Elder and Lubotsky (2009); Evans et al. (2010); Mühlenweg et al. (2012)

Appendix

Related literature

But also non-cognitive benefits:

- Less likely to be classified with ADHD
  $\longrightarrow$ Dhuey and Lipscomb (2010); Elder and Lubotsky (2009); Evans et al. (2010); Mühlenweg et al. (2012)
- More persistent and less irritable
  $\longrightarrow$ Mühlenweg et al. (2012)

Appendix

Related literature

But also non-cognitive benefits:

- Less likely to be classified with ADHD
  $\longrightarrow$ Dhuey and Lipscomb (2010); Elder and Lubotsky (2009); Evans et al. (2010); Mühlenweg et al. (2012)
- More persistent and less irritable
  $\longrightarrow$ Mühlenweg et al. (2012)
- Less likely to suffer from bullying
  $\longrightarrow$ Mühlenweg and Puhani (2010)

Appendix

Related literature

But also non-cognitive benefits:

- Less likely to be classified with ADHD
  $\longrightarrow$ Dhuey and Lipscomb (2010); Elder and Lubotsky (2009); Evans et al. (2010); Mühlenweg et al. (2012)
- More persistent and less irritable
  $\longrightarrow$ Mühlenweg et al. (2012)
- Less likely to suffer from bullying
  $\longrightarrow$ Mühlenweg and Puhani (2010)
- More likely to to hold leadership positions as teenagers
  $\longrightarrow$ Dhuey and Lipscomb (2008)

Appendix
Related literature

But also non-cognitive benefits:

- Less likely to be classified with ADHD
  $\longrightarrow$ Dhuey and Lipscomb (2010); Elder and Lubotsky (2009); Evans et al. (2010); Mühlenweg et al. (2012)
- More persistent and less irritable
  $\longrightarrow$ Mühlenweg et al. (2012)
- Less likely to suffer from bullying
  $\longrightarrow$ Mühlenweg and Puhani (2010)
- More likely to to hold leadership positions as teenagers
  $\longrightarrow$ Dhuey and Lipscomb (2008)
- Less likely to commit crimes or be incarcerated
  $\longrightarrow$ Landersø et al. (2017); Cook and Kang (2016); Dhuey et al. (2017)

Appendix

Related literature ▸ Back

Impacts on adult outcomes are more ambiguous:

Appendix

Related literature ▸ Back

Impacts on adult outcomes are more ambiguous:

- Link to higher wages later in working career
  $\longrightarrow$ Fredriksson and Öckert (2014)

Related literature

▸ Back

Impacts on adult outcomes are more ambiguous:

- Link to higher wages later in working career
  $\longrightarrow$ Fredriksson and Öckert (2014)
- More likely to become a corporate CEO
  $\longrightarrow$ Du et al. (2012)

Appendix

Related literature

▸ Back

Impacts on adult outcomes are more ambiguous:

- Link to higher wages later in working career
  $\longrightarrow$ Fredriksson and Öckert (2014)
- More likely to become a corporate CEO
  $\longrightarrow$ Du et al. (2012)
- No significant effects on prime-age earnings
  $\longrightarrow$ Black et al. (2011); Dobkin and Ferreira (2010); Fredriksson and Öckert (2014)

Appendix

Local polynomial specifications

Intent to treat effects (ITT)

$$\min \sum_{i=1}^{N(h)} \left(Y_{ig} - \alpha_0 - \alpha\tau_i - \underbrace{\mathbf{f}(B_i)}_{\text{Trend}} - \mathbf{X}_i\delta - \underbrace{\varphi_c}_{\text{Cohort FE}} \right)^2 \underbrace{\mathbf{K}_h(\tau_i, B_i)}_{\triangle \text{ kernel}}$$

Appendix

Local polynomial specifications

Intent to treat effects (ITT)

$$\min \sum_{i=1}^{N(h)} \left(Y_{ig} - \alpha_0 - \alpha\tau_i - \underbrace{\mathbf{f}(B_i)}_{\text{Trend}} - \mathbf{X}_i\delta - \underbrace{\varphi_c}_{\text{Cohort FE}} \right)^2 \underbrace{\mathbf{K}_h(\tau_i, B_i)}_{\triangle\ \text{kernel}}$$

$$\mathbf{f}(B_i) = \phi_1 B_i + \phi_2 \tau_i B_i$$

Appendix

Local polynomial specifications

Intent to treat effects (ITT)

$$\min \sum_{i=1}^{N(h)} \Bigg(Y_{ig} - \alpha_0 - \alpha\tau_i - \underbrace{\mathbf{f}(B_i)}_{\text{Trend}} - \mathbf{X}_i\delta - \underbrace{\varphi_c}_{\text{Cohort FE}} \Bigg)^2 \underbrace{\mathbf{K}_h(\tau_i, B_i)}_{\triangle \text{ kernel}}$$

$$\mathbf{f}(B_i) = \phi_1 B_i + \phi_2 \tau_i B_i \quad \textbf{or} \quad \mathbf{f}(B_i) = \sum_{p=1}^{2} \phi_p B_i^p + \sum_{p=1}^{2} \phi_{pp} \tau_i B_i^p$$

Appendix

Local polynomial specifications

Intent to treat effects (ITT)

$$\min \sum_{i=1}^{N(h)} \left(Y_{ig} - \alpha_0 - \alpha\tau_i - \underbrace{\mathbf{f}(B_i)}_{\text{Trend}} - \mathbf{X}_i\delta - \underbrace{\varphi_c}_{\text{Cohort FE}} \right)^2 \underbrace{\mathbf{K}_h(\tau_i, B_i)}_{\triangle\ \text{kernel}}$$

$$\mathbf{f}(B_i) = \phi_1 B_i + \phi_2 \tau_i B_i \quad \textbf{or} \quad \mathbf{f}(B_i) = \sum_{p=1}^{2} \phi_p B_i^p + \sum_{p=1}^{2} \phi_{pp} \tau_i B_i^p$$

$$\mathbf{K}_h(\tau_i, B_i) = \max\left(0, 1 - \left|\frac{B_i}{h}\right|\right)$$

Local average treatment effects (LATE)

$$\min \sum_{i=1}^{N(h)} (A_i - \theta_0 - \theta\tau_i - \mathbf{f}(B_i) - \mathbf{X}_i\delta - \varphi_c)^2 \mathbf{K}_h(\tau_i, B_i)$$

Local average treatment effects (LATE)

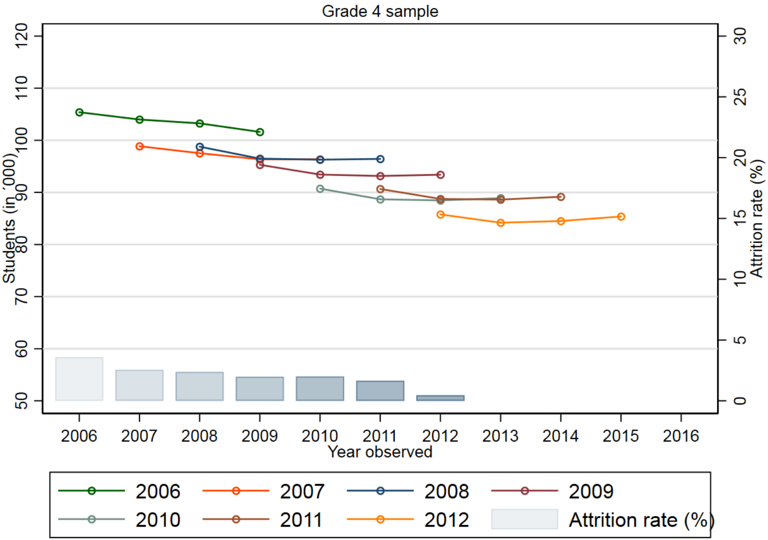
$$\min \sum_{i=1}^{N(h)} \left(A_i - \theta_0 - \theta\tau_i - \mathbf{f}(B_i) - \mathbf{X}_i\delta - \varphi_c\right)^2 \mathbf{K}_h(\tau_i, B_i)$$

$$\min \sum_{i=1}^{N(h)} \left(Y_{ig} - \beta_0 - \beta\hat{A}_i - \mathbf{f}(B_i) - \mathbf{X}_i\delta - \varphi_c\right)^2 \mathbf{K}_h(\tau_i, B_i)$$

Appendix

Attrition Grade 4 sample

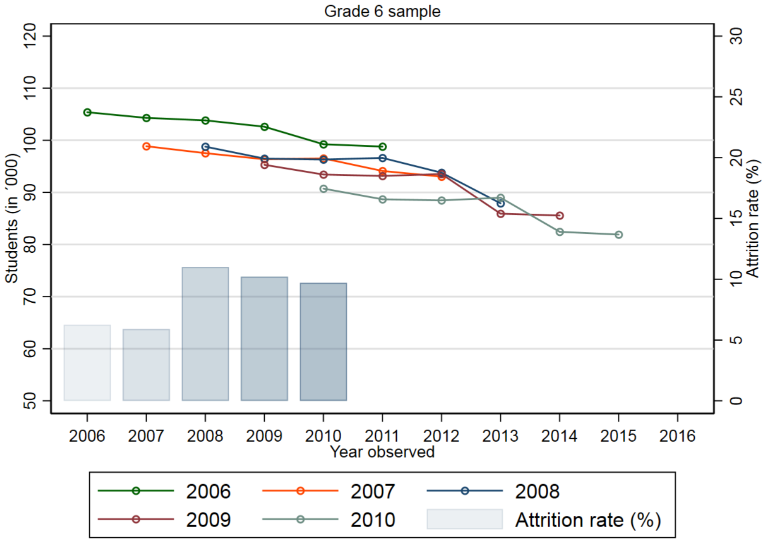
▸ Back



Appendix

Attrition Grade 6 sample

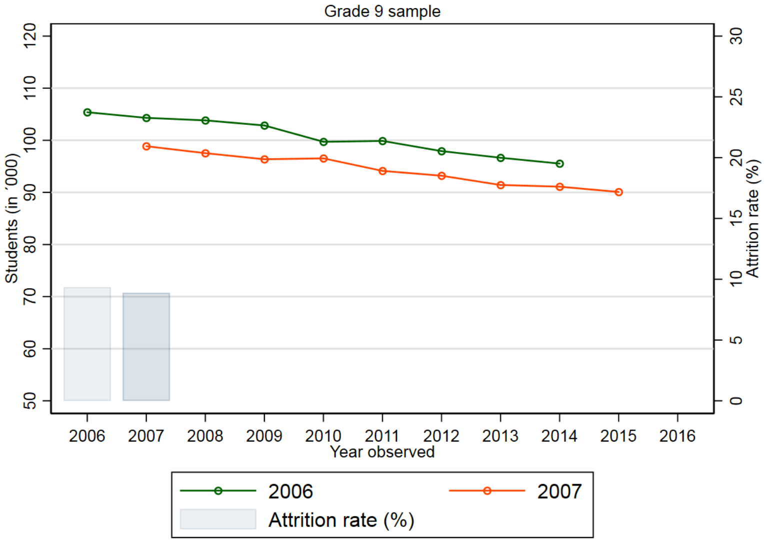
▸ Back



Appendix

Attrition Grade 9 sample

‣ Back



Appendix

Descriptive statistics

Back

Sample:	Full sample		60-days before cutoff		60-days after cutoff		Difference
	Obs.	%	Obs.	%	Obs.	%	p-value
Student characteristics							
Grade 4 sample							
Female	229,661	48.61	36,220	48.56	37,145	48.40	0.67
First generation immigrant	229,661	2.31	36,220	2.26	37,145	2.27	0.92
Access to computer at home	229,661	55.39	36,220	55.10	37,145	55.15	0.89
School social support (ASE)	229,661	38.76	36,220	39.18	37,145	39.70	0.15
Dad unemployed	229,661	6.93	36,220	7.16	37,145	7.05	0.61
Household with higher education	229,661	21.67	36,220	21.10	37,145	21.57	0.15
Grade 6 sample							
Female	300,182	49.37	46,198	49.96	49,066	49.17	0.01
First generation immigrant	300,182	2.49	46,198	2.33	49,066	2.58	0.01
Access to computer at home	300,182	46.09	46,198	45.63	49,066	46.40	0.02
School social support (ASE)	300,182	23.77	46,198	23.73	49,066	24.10	0.19
Dad unemployed	300,182	4.66	46,198	4.87	49,066	4.63	0.12
Household with higher education	300,182	17.85	46,198	17.69	49,066	17.85	0.56
Grade 9 sample							
Female	188,648	51.51	28,676	52.21	31,125	51.38	0.04
First generation immigrant	188,648	2.27	28,676	2.07	31,125	2.53	0.00
Access to computer at home	188,648	47.91	28,676	47.41	31,125	48.32	0.03
School social support (ASE)	188,648	16.24	28,676	16.28	31,125	16.33	0.87
Dad unemployed	188,648	3.67	28,676	3.64	31,125	3.72	0.63
Household with higher education	188,648	20.04	28,676	20.08	31,125	19.73	0.32

Appendix

Continuity in covariates

Back

Grade sample:	Grade 4	Grade 6	Grade 9	Controls
Outcome:	Coef. (SE)	Coef. (SE)	Coef. (SE)	
Female	-0.016 (0.022)	0.004 (0.017)	0.019 (0.022)	No
	-0.016 (0.022)	0.005 (0.017)	0.020 (0.022)	Yes
Immigrant	0.015 (0.008)	0.007 (0.006)	0.008 (0.008)	No
	0.015 (0.008)	0.007 (0.006)	0.008 (0.008)	Yes
School social support	0.059 (0.020)	-0.012 (0.011)	-0.024 (0.012)	No
	0.057 (0.018)	-0.014 (0.011)	-0.027 (0.011)	Yes
Unemployed dad	-0.008 (0.013)	0.002 (0.005)	0.009 (0.010)	No
	-0.012 (0.013)	0.003 (0.005)	0.010 (0.010)	Yes
Computer at home	-0.014 (0.016)	0.004 (0.021)	0.018 (0.019)	No
	-0.006 (0.015)	0.005 (0.020)	0.019 (0.019)	Yes
Higher education in HH	-0.007 (0.018)	-0.008 (0.014)	-0.009 (0.012)	No
	0.007 (0.017)	-0.010 (0.013)	-0.015 (0.011)	Yes
Observations	72807	94573	59345	

Appendix

First-stage estimates

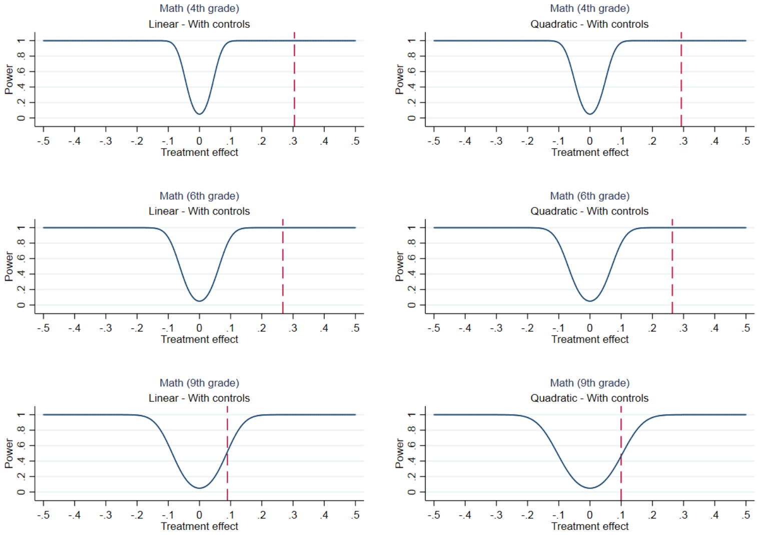
Back

Outcome: school starting age	30-days bandwidth		60-days bandwidth		MSE-optimal bandwidth	
	(1)	(2)	(3)	(4)	(5)	(6)
Grade 4						
τ_4	0.674	0.670	0.709	0.706	0.686	0.685
	(0.038)	(0.038)	(0.026)	(0.026)	(0.013)	(0.014)
Observations	36124	36124	72807	72807	113709	108656
Bandwidth (in days)	30	30	60	60	93	89
Grade 6						
τ_6	0.730	0.729	0.740	0.739	0.733	0.731
	(0.014)	(0.014)	(0.010)	(0.011)	(0.008)	(0.008)
Observations	47125	47125	94573	94573	105815	104133
Bandwidth (in days)	30	30	60	60	66	65
Grade 9						
τ_9	0.737	0.736	0.748	0.746	0.742	0.741
	(0.023)	(0.023)	(0.015)	(0.016)	(0.010)	(0.010)
Observations	29428	29428	59345	59345	84777	82809
Bandwidth (in days)	30	30	60	60	84	83
Polynomial order	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic
Student controls	NO	YES	NO	YES	NO	YES
Cohort FEs	NO	YES	NO	YES	NO	YES

Appendix

Statistical power

Back



Appendix

Only non-repeaters

Back

Outcome:	Math				Language			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grade 4								
School starting age	0.251	0.280	0.245	0.275	0.342	0.370	0.325	0.356
	(0.029)	(0.025)	(0.034)	(0.029)	(0.029)	(0.022)	(0.051)	(0.045)
Observations	31138	30064	45349	44192	63021	78135	57492	60766
Bandwidth (in days)	29	28	41	41	57	72	53	56
Grade 6								
School starting age	0.195	0.218	0.179	0.206	0.235	0.255	0.242	0.261
	(0.031)	(0.024)	(0.035)	(0.026)	(0.029)	(0.025)	(0.031)	(0.027)
Observations	58094	56681	67816	73176	59572	60964	86475	86475
Bandwidth (in days)	43	42	49	54	43	45	64	64
Grade 9								
School starting age	0.073	0.105	0.150	0.152	0.170	0.182	0.180	0.187
	(0.035)	(0.032)	(0.045)	(0.039)	(0.032)	(0.030)	(0.034)	(0.031)
Observations	38128	32081	32991	36401	34910	40944	51910	71246
Bandwidth (in days)	44	38	38	43	41	47	61	83
Polynomial order	Linear	Linear	Quadratic	Quadratic	Linear	Linear	Quadratic	Quadratic
Student controls	NO	YES	NO	YES	NO	YES	NO	YES
Cohort Fes	NO	YES	YES	YES	NO	YES	YES	YES

Appendix

With birth day of the week FEs

Back

Outcome:	Math performance (1)	(2)	Language performance (3)	(4)	Grade retention (5)	(6)
(Until) Grade 4						
School starting age	0.289	0.274	0.364	0.360	-0.060	-0.054
	(0.022)	(0.024)	(0.034)	(0.040)	(0.009)	(0.011)
Observations	34866	53287	41015	65624	57032	66960
Bandwidth (in days)	29	44	34	53	47	54
(Until) Grade 6						
School starting age	0.262	0.255	0.294	0.291	-0.081	-0.079
	(0.025)	(0.028)	(0.023)	(0.024)	(0.010)	(0.011)
Observations	64417	100530	61354	115385	61467	93673
Bandwidth (in days)	41	64	38	72	39	58
(Until) Grade 9						
School starting age	0.095	0.150	0.192	0.179	-0.047	-0.048
	(0.024)	(0.038)	(0.024)	(0.032)	(0.009)	(0.010)
Observations	56529	39324	66312	69474	43488	76707
Bandwidth (in days)	57	39	66	70	43	76
Polynomial order	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic

Appendix

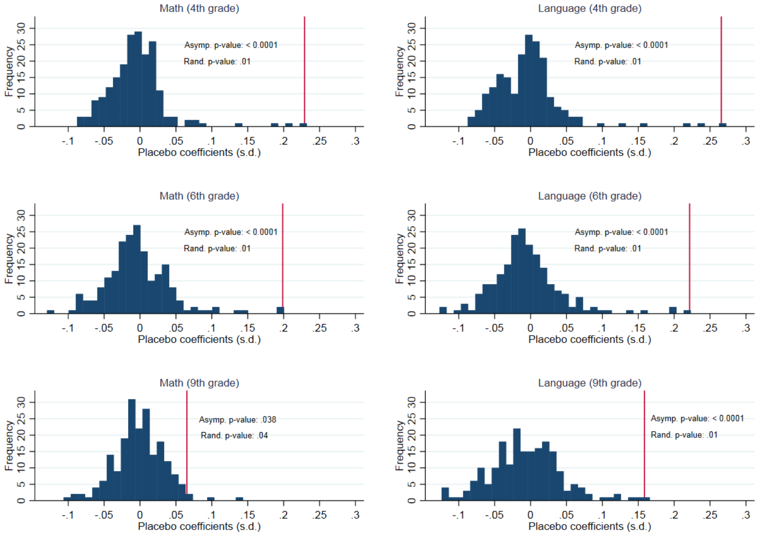
With birth day of the week FEs ▸ Back

Outcome:	Math performance			Language performance			Grade retention		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Until) Grade 4									
School starting age	0.288	0.272	0.274	0.393	0.356	0.373	-0.041	-0.052	-0.053
	(0.028)	(0.024)	(0.024)	(0.075)	(0.048)	(0.035)	(0.017)	(0.013)	(0.011)
Observations	35742	72001	61461	35767	72040	71283	35826	72155	68302
Left bandwidth (in days)	30	60	53	30	60	66	30	60	54
Right bandwidth (in days)	30	60	47	30	60	50	30	60	64
(Until) Grade 6									
School starting age	0.194	0.239	0.247	0.326	0.302	0.291	-0.071	-0.079	-0.078
	(0.042)	(0.030)	(0.028)	(0.031)	(0.031)	(0.024)	(0.014)	(0.013)	(0.011)
Observations	46958	94251	87299	47043	94393	123601	47125	94573	101436
Left bandwidth (in days)	30	60	63	30	60	73	30	60	57
Right bandwidth (in days)	30	60	48	30	60	82	30	60	69
(Until) Grade 9									
School starting age	0.168	0.155	0.153	0.152	0.184	0.175	-0.032	-0.036	-0.039
	(0.060)	(0.038)	(0.037)	(0.049)	(0.036)	(0.031)	(0.013)	(0.011)	(0.010)
Observations	29245	58995	49042	29407	59290	65373	29428	59345	54603
Left bandwidth (in days)	30	60	42	30	60	70	30	60	49
Right bandwidth (in days)	30	60	56	30	60	61	30	60	61

Appendix

Robustness checks → Randomization-based p-values

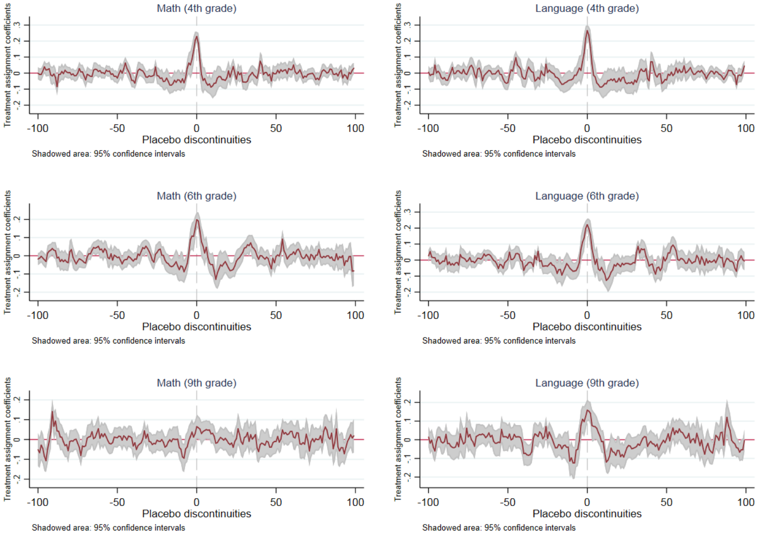
Back



Appendix

Robustness checks → Placebo coefficients

Back



Appendix

Mechanisms → Age-at-test vs. school starting age effects ‣ Back

Why is there a decline in SSA effects?

Why is there a decline in SSA effects?

- SSA effect orthogonal to age-at-test effects may be negative
  → Peña (2017); Black et al. (2011); Crawford et al. (2007)

Why is there a decline in SSA effects?

- SSA effect orthogonal to age-at-test effects may be negative
  → Peña (2017); Black et al. (2011); Crawford et al. (2007)

$$\underbrace{\frac{\partial h_t}{\partial A}}_{\text{Absolute age effect}}$$

Why is there a decline in SSA effects?

- SSA effect orthogonal to age-at-test effects may be negative
  → Peña (2017); Black et al. (2011); Crawford et al. (2007)

$$\underbrace{\frac{\partial h_t}{\partial A}}_{\text{Absolute age effect}} = \underbrace{\frac{\partial h_t}{\partial t}}_{\text{Age-at-test effect}}$$

Why is there a decline in SSA effects?

- SSA effect orthogonal to age-at-test effects may be negative
  → Peña (2017); Black et al. (2011); Crawford et al. (2007)

$$\underbrace{\frac{\partial h_t}{\partial A}}_{\text{Absolute age effect}} = \underbrace{\frac{\partial h_t}{\partial t}}_{\text{Age-at-test effect}} + \underbrace{\frac{\partial h_t}{\partial SSA_\perp}}_{\perp\ \text{SSA effect}}$$

Appendix

Mechanisms → Minimal assumptions model

▸ Back

$$\frac{\partial h_t}{\partial t} \geq 0 \quad \wedge$$

Appendix

Mechanisms → Minimal assumptions model

Back

$$\frac{\partial h_t}{\partial t} \geq 0 \quad \wedge \quad \frac{\partial^2 h_t}{\partial t^2} \leq 0$$

Appendix

Mechanisms → Minimal assumptions model

Back

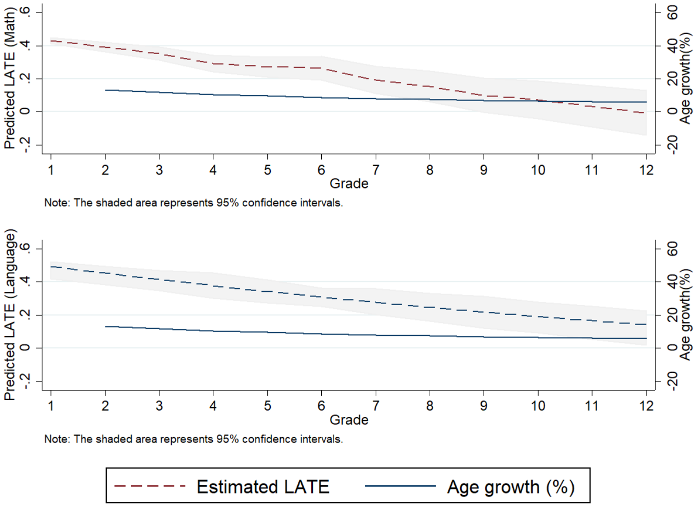
$$\frac{\partial h_t}{\partial t} \geq 0 \quad \wedge \quad \frac{\partial^2 h_t}{\partial t^2} \leq 0$$

$$\frac{\partial h_t}{\partial t} \propto \underbrace{g(A_t) = \frac{A_t - A_{t-1}}{A_{t-1}} - 1}_{\text{Growth in age}}$$

Appendix

Mechanisms → Rates of decline

▸ Back



Appendix

Mechanisms → Rates of decline ▸ Back

Given our empirical results:

$$\underbrace{\frac{\partial^2 h_t}{\partial A^2} < \frac{\partial g(A_t)}{\partial t}}_{\text{Estimated LATE fall quicker}}$$

Appendix

Mechanisms $\rightarrow$ Rates of decline

Back

Given our empirical results:

$$\underbrace{\frac{\partial^2 h_t}{\partial A^2} < \frac{\partial g(A_t)}{\partial t}}_{\text{Estimated LATE fall quicker}} \implies \frac{\partial^2 h_t}{\partial A^2} < \frac{\partial^2 h_t}{\partial t^2}$$

Appendix

Mechanisms → Rates of decline

Back

Given our empirical results:

$$\underbrace{\frac{\partial^2 h_t}{\partial A^2} < \frac{\partial g\left(A_t\right)}{\partial t}}_{\text{Estimated LATE fall quicker}} \implies \frac{\partial^2 h_t}{\partial A^2} < \frac{\partial^2 h_t}{\partial t^2} \implies \frac{\partial h_t}{\partial SSA_\perp} < 0$$