

Early tracking and achievement growth

Maciej Jakubowski
Directorate for Education, OECD

December 2009

Outline

- Definition and importance of „early tracking“
- Existing evidence
- Data and methods used in this study
- DiD Country-level approach
- DiD regression on individual data adjusting for discrepancies in background characteristics
- DiD Quantile regression – comparing the impact of tracking for low- and high-achievers
- Discussion

What is early tracking?

- Separate schools with different programmes
- In this study, early tracking systems are those which separate students younger than 15-year-old
 - Tracking countries: Germany or Czech Republic
 - Non-tracking countries: Sweden or England
- Institutional tracking is one of the major characteristics of any educational system
- It probably heavily affects how students progress through the school system on their way to labour market
- Changing the timing of tracking is costly and politically very difficult

Existing evidence

International studies provide mixed evidence:

- Hanushek and Woessmann (2006): tracking increases inequalities with no positive effect on average performance; Ammermueller (2005): similar findings
- Waldinger (2007) and Jakubowski (2010): these results are not robust, heavily depend on the applied methodology
- Brunello and Checchi (2007): IALS, positive on achievement, negative on income

Country studies support comprehensive schools:

- Finland - Pekkarinen, Uusitalo, Kerr (2009): small positive impact
- Germany - Mühlenweg (2007) : positive for low achievers
- Poland - Jakubowski, Patrinos, Porta, Wisniewski (forthcoming): introduction of comprehensive lower-secondary schools for 15-year-olds boosted achievement of those previously in vocational schools
- Poland – Jakubowski, Pokropek (forthcoming): These benefits partly disappear in upper secondary schools with vocational schools negatively affecting student achievement growth

Experiments support tracking

- Kenya - Duflo, Dupas, and Kremer, 2008: tracking have longstanding positive effects for students at different ability levels

Data

Existing comparisons of student achievement in primary and secondary schools

Domain	Primary school	Secondary school
Reading	PIRLS 2001	PISA 2000
Mathematics	TIMSS 2003	PISA 2003
Science	TIMSS 2007	PISA 2006

Control and outcome variables

□ Control variables recoded to similar categories in PISA, PIRLS and TIMSS:

- Gender
- Immigrant background (born outside the country, speaking language of the test)
- Number of books at home
- Grade
- Age

□ Outcome:

- Plausible values in main domains rescaled to 500/100 in the sample of analyzed countries
- Senate weights used to equally weight countries
- Simplified analysis : 1st plausible value + analytical and bootstrapped SE

Difference-in-differences method

$$\text{DiD} = (\textit{post} - \textit{pre} / \textit{treated}) \\ - (\textit{post} - \textit{pre} / \textit{control})$$

- Control group: non-tracking countries
- Treatment group: tracking countries
- Pre: primary school (PIRLS, TIMSS)
- Post: secondary school (PISA)

NOTE: No tracking in primary schools!

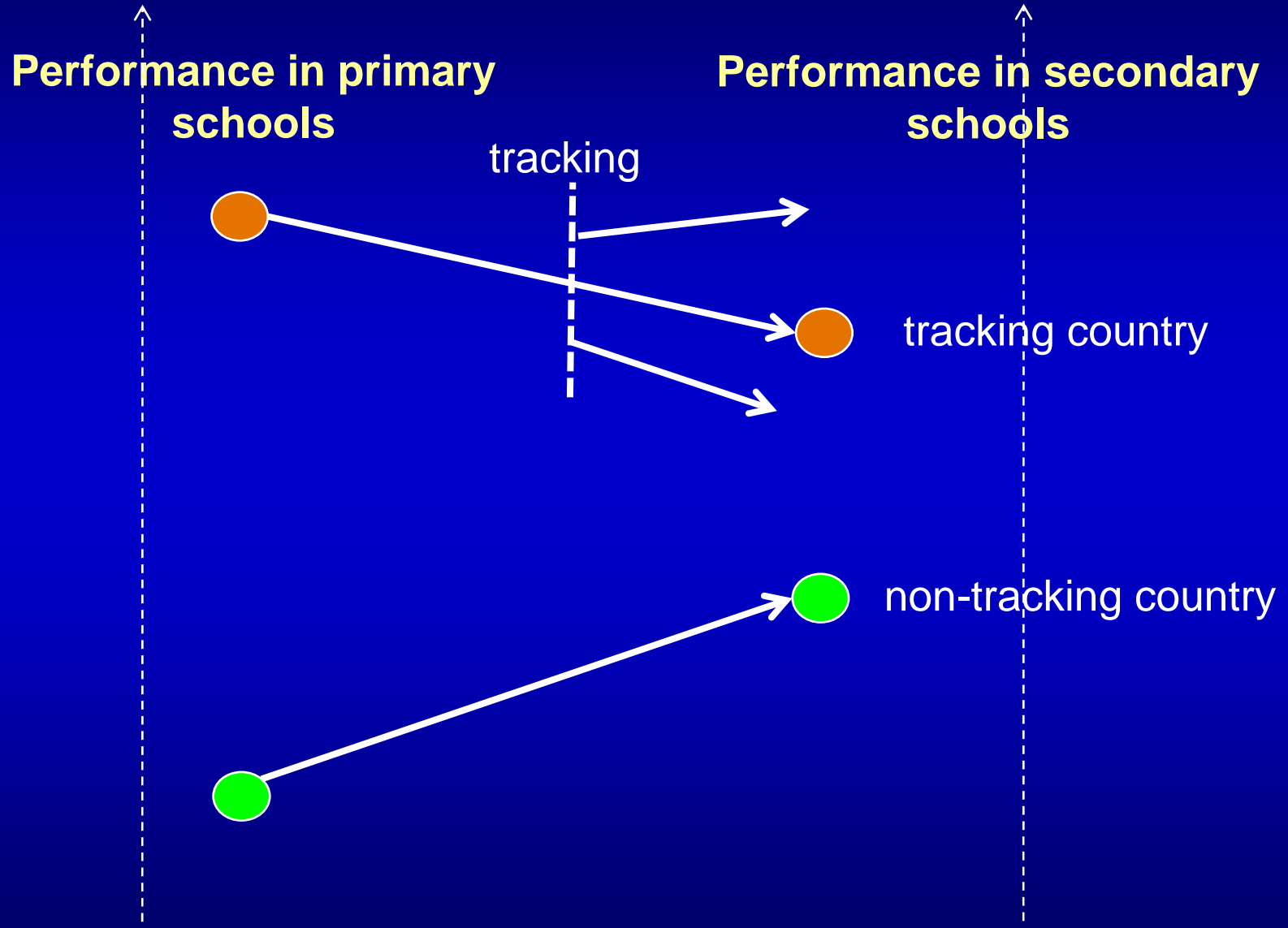
Performance in secondary schools



tracking country



non-tracking country



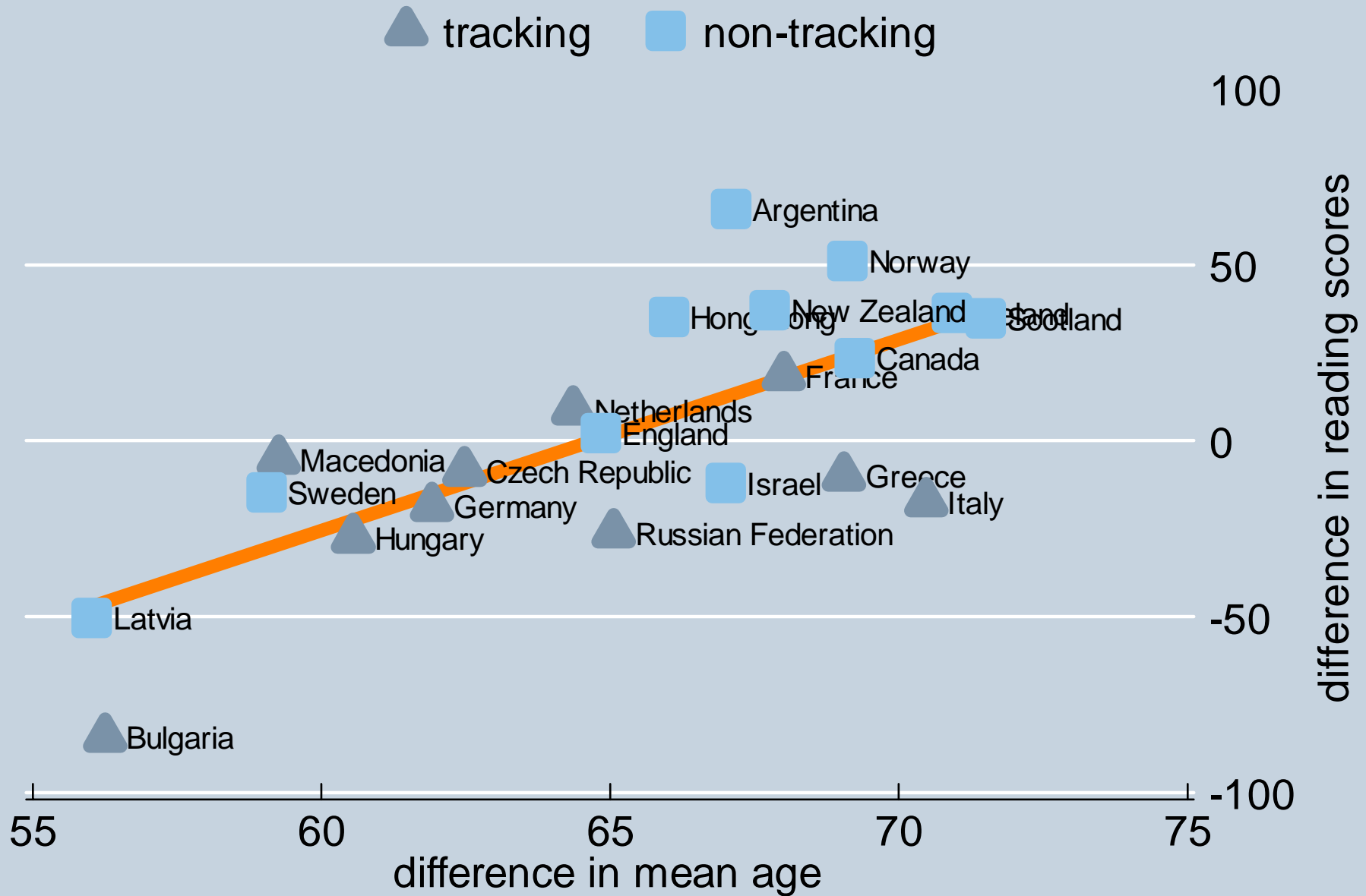
Different ways of applying this approach:

- 1) Country-level regression with secondary school mean achievement regressed on prior achievement (Hanushek & Woessmann, 2006)
- 2) Regression on the pooled individual data: tracking effects on the mean performance + the differential effect for disadvantage students
- 3) Quantile DD regression: quantile effects of tracking for low and high achieving students

Robustness checks of country level approach

- **Jakubowski (2010):**
 - PISA and PIRLS/TIMSS are not directly comparable because they are representative to different populations (age vs. grade)
 - Adjustments using micro data, re-estimation using country-level regression
 - Controlling for mean age differences across countries/surveys
 - Making samples more comparable: modal grades, different age brackets, with/without migrant students
 - Testing effects for different domains/survey combinations
- **Hanushek and Woessmann results are not robust**
- **Many results non-significant, some positive, some negative**

PISA and PIRLS: age and performance



Different method: DiD regression on the pooled individual datasets

- Directly controlling for differences in individual characteristics and their change over time: age, grade, gender, number of books at home, migrant status
- country fixed effects
- Testing whether tracking effects are similar for low- and high-SES students (DiDiD approach)
- Expectation: we should observe negative effect for low-SES students
- Testing whether tracking effects are similar in early and late tracking countries or Western/Eastern European countries

reading literacy in PIRLS/PISA testing heterogeneity of tracking effects

Early tracking*Time	-2.26	-7.31**
Late tracking*Time	-27.49***	-16.32***
Low SES background*Early tracking*Time	-3.39	-2.26
Low SES background*Late tracking*Time	2.66	1.96
Eastern European country		67.67***
Eastern European country*time		-25.79***

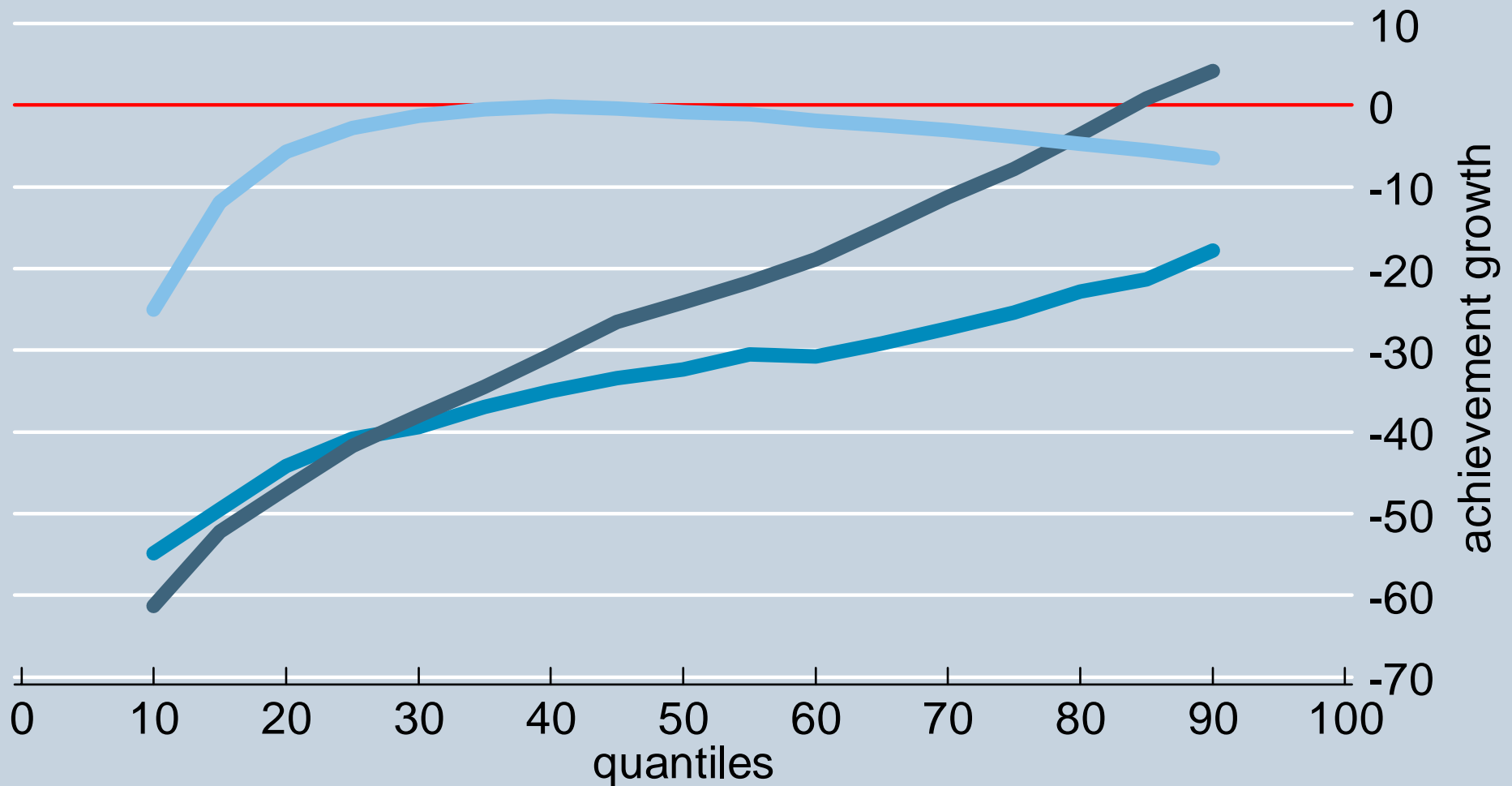
Quantile effects

- Tracking effects estimated for different achievement levels: from the 10th to the 90th percentile of student achievement distribution
- Uses individual data, but adjusts for age effects only not to confound the impact of tracking with anything else (e.g. retention policy)
- Shows how tracking affects students depending on their achievement level
- Expectation: low achieving students should be affected negatively, while the effect for high-achievers should be positive or at least non-negative

Tracking effects by domain

age adjusted

Reading Math Science

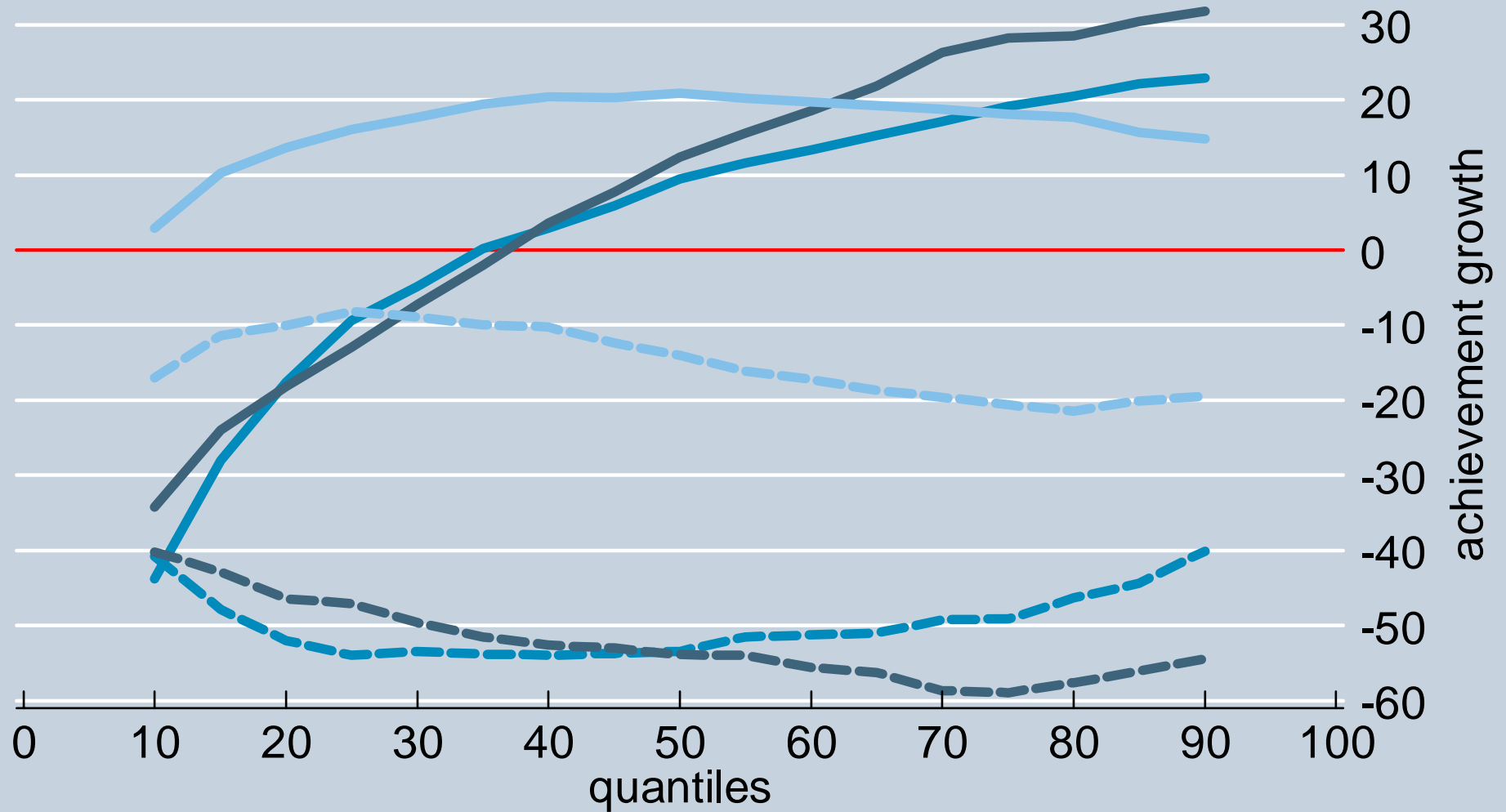


Is that tracking or something else?

- These results suggest that tracking has negative impact on low achievers in reading and math
- However, achievement growth for high-achievers is also smaller in tracking countries
- Is tracking confounded with some other policies/characteristics of tracking countries?
- Some tracking countries are Eastern European countries (e.g. Hungary), while some are tracking students much earlier than others (e.g. Germany)

Early trackers and Eastern Europe

Early tracking Reading Math Science
Eastern Europe Reading Math Science



Conclusions

- PIRLS and TIMSS data have to be adjusted for age differences before any meaningful comparisons with PISA could be made
- Country level approach is not robust even to minor modifications
- However, using micro data we partly confirmed the negative impact of tracking on student achievement growth
- Early tracking negatively affects performance of low-achieving students in reading and math, while it increases performance of high-achievers
- However, there is no evidence on the negative impact of tracking in science
- Impact of tracking has to be separated from overall lower achievement growth in Eastern European countries