American Economic Review: Papers & Proceedings 2014, 104(5): 400–405
http://dx.doi.org/10.1257/aer.104.5.400

POLICY INTERVENTIONS AND EDUCATIONAL OUTCOMES‡

Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement†

By Kalena E. Cortes and Joshua S. Goodman*

One of the most controversial educational practices among American public schools is the use of *ability-tracking* of students. Over the years, tracking has received harsh criticism. In fact, policymakers have launched initiatives to discourage tracking across the nation, and started recommending districts place students with different ability levels in the same classrooms (Wheelock 1992). Opponents argue that tracking not only fails to benefit students, but in racially integrated schools, low-track classes have an overrepresentation of minority students. Supporters of the gifted and talented programs maintain that heterogeneous classes inhibit bright students from learning, and that bright students only languish in mixed ability classes.

To date, the research on tracking has been quite mixed and inconclusive (see, for example, Betts and Shkolnik 2000; Figlio and Page 2002; Clark 2010).[1] Recent papers on elite exam schools in the United States (Abdulkadiroglu, Angrist, and Pathak 2011; Dobbie and Fryer 2011) and Kenya (Lucas and Mbiti forthcoming), as well as gifted and talented programs (Bui, Craig, and Imberman 2011), all find surprisingly few positive impacts of being exposed to a very different set of peers generated by discontinuities in admissions processes. The few exceptions to this literature are studies by Duflo, Dupas, and Kremer (2011), who find that tracking benefits both lower- and higher-ability students in Kenya; and Pop-Eleches and Urquiola (2013) who find a modest but significant impact on end-of-high school exam scores associated with entry to the baccalaureate track in Romanian high schools.

This paper provides new evidence on tracking by studying an innovative curriculum implemented in Chicago Public Schools (CPS). In 2003, CPS enacted a new algebra policy, which required ninth grade students with incoming eighth grade math test scores that were below the national median to take two periods of algebra for a full year—regular algebra plus an algebra with support class. Also, teachers who taught algebra under this new policy were provided curricular resources and professional development to help them effectively use the extra instructional time. The “double-dose” algebra policy led CPS schools to *sort* students into algebra classes by the student’s eighth grade math ability. Thus, as a consequence of this policy, tracking increased in all algebra classes.

‡*Discussants:* Monica Deza, University of Texas-Dallas; Damon Jones, University of Chicago; Juan Carlos Suárez Serrato, Stanford University; Omari Swinton, Howard University.

* Cortes: Texas A&M University, 4220 TAMU, 1049 Allen Building, College Station, TX 77843 (e-mail: kcortes@tamu.edu); Goodman: Harvard University, Taubman 354, 79 John F. Kennedy St., Cambridge, MA 02138 (e-mail: joshua_goodman@hks.harvard.edu). This research was conducted while Cortes was a Visiting Scholar at Stanford’s Graduate School of Education, Center for Education Policy Analysis, 2013–2014; kcortes@stanford.edu. This research was funded by the United States Department of Education, the Institute of Education Sciences under award R305A120466. Institutional support from Texas A&M University, Stanford’s Center for Education Policy Analysis, and Harvard’s Taubman Center for State and Local Government are also gratefully acknowledged.

† Go to http://dx.doi.org/10.1257/aer.104.5.400 to visit the article page for additional materials and author disclosure statement(s).

[1] For a thoughtful and detailed review of the earlier literature on tracking, see Betts (2011).

Using a longitudinal dataset from CPS that follows students from eighth grade through high school graduation, we implement a difference-in-difference strategy to identify the impacts of this policy on student academic performance. We show that double-dosed students are exposed to a much lower-skilled group of peers in their algebra classes but nonetheless benefit substantially from the additional instructional time and improved pedagogy. Specifically, we find that the policy increased freshman academic performance, and most importantly, had positive longer-term effects on later math coursework and test scores.

I. Data and Empirical Strategy

We use transcript data from CPS that follow students from eighth grade through high school. These data include demographics, detailed high school transcripts, and standardized test scores. Our sample consists of all regular education students entering ninth grade for the first time in the fall of 2001–2004 who were enrolled in at least one algebra class. The first two cohorts, 2001 and 2002, are untreated and the second two cohorts, 2003 and 2004, are treated. The main independent variable, which will provide our instrument, is each student's eighth grade score on the math portion of the Iowa Test of Basic Skills. All CPS eighth graders are required to take this test. As shown in Cortes, Goodman, and Nomi (2013a, 2013b), though 48 percent of CPS students score below the fiftieth percentile and are double-dose eligible, only 43 percent enroll in double-dose algebra, suggesting imperfect compliance with the rule.

Comparison of the outcomes of double-dosed and non-double-dosed students might yield biased estimates of the program's impacts given potentially large differences in unobserved characteristics between the two groups of students. To eliminate this potential bias, we exploit the fact that students scoring below the fiftieth percentile on the eighth grade math test were supposed to enroll in double-dose algebra. This allows us to use a difference-in-differences (DID) approach to identify the impact of the policy, using the assignment rule as an exogenous source of variation in the probability that a given student will be double-dosed. We implement the DID framework using the regressions below:

$$
\begin{aligned}
(1) \qquad Y_{ist} &= \alpha_0 + \alpha_1 \cdot LowScore_{ist} \\
&\quad + \alpha_2 \cdot Post_{ist} \\
&\quad + \alpha_3 \cdot LowScore_{ist} \cdot Post_{ist} \\
&\quad + X_{ist} \cdot \varphi + \eta_{ist}
\end{aligned}
$$

$$
\begin{aligned}
(2) \quad DoubleDose_{ist} &= \gamma_0 + \gamma_1 \cdot LowScore_{ist} \\
&\quad + \gamma_2 \cdot Post_{ist} \\
&\quad + \gamma_3 \cdot LowScore_{ist} \cdot Post_{ist} \\
&\quad + X_{ist} \cdot \varphi + \mu_{ist}
\end{aligned}
$$

$$
\begin{aligned}
(3) \qquad Y_{ist} &= \beta_0 + \beta_1 \cdot LowScore_{ist} \\
&\quad + \beta_2 \cdot Post_{ist} \\
&\quad + \beta_3 \cdot DoubleDose_{ist} \\
&\quad + X_{ist} \cdot \varphi + \varepsilon_{ist},
\end{aligned}
$$

where, Y_{ist} represents the outcome of interest for student i in high school s in cohort t, $LowScore_{ist}$ indicates an eighth grade math score below the fiftieth percentile, $Post_{ist}$ indicates the 2003 and 2004 treated cohorts, $DoubleDose_{ist}$ is a double-dose indicator, and X_{ist} is a vector of student demographics, neighborhood characteristics, and high school fixed effects. X_{ist} includes gender, race/ethnicity, eighth grade reading score, free and reduced price lunch status, high school start age, cohort, socioeconomic, and poverty measures constructed for each student's residential block group from the 2000 census.

By controlling for differences between low- and high-scoring students in the 2001 and 2002 pretreatment cohorts and for overall differences between cohorts, the interaction coefficient (α_3) from equation (1) estimates how the difference in outcomes between low- and high-scoring students changed at the time double-dose algebra was introduced. This reduced form equation produces an intention-to-treat estimate because compliance with the assignment rule was imperfect. We therefore use equation (2) as a first-stage to predict how introduction of the policy affected the probability of a low-scoring student being double-dosed. Our ultimate estimate of interest is therefore the double-dose coefficient (β_3) from equation (3), in which

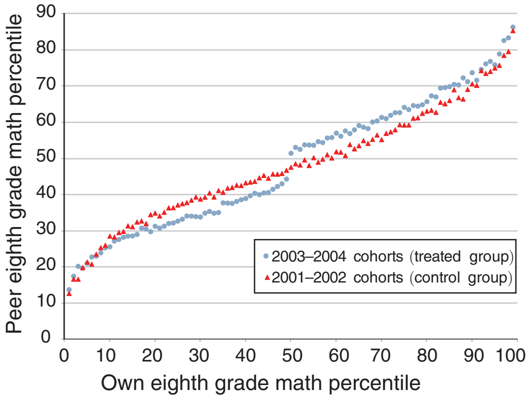

Figure 1. The Double-Dose Policy and Peer Composition in Freshman Algebra

Notes: Shown above is the mean eighth grade math percentile of each student's ninth grade regular algebra class peers, averaged by the student's own eighth grade math percentile. The 2003–2004 cohorts were subject to the double-dose algebra policy while the 2001–2002 cohorts were not.

Source: Authors' calculations from Chicago Public Schools High School Transcripts Data, 2001–2004 student cohorts.

$DoubleDose_{ist}$ has been instrumented with the interaction of $LowScore_{ist}$ and $Post_{ist}$ using equation (2) as our first-stage. Our first-stage results (not reported here) show that low-scoring students were 70 percentage points more likely to be double-dosed than high-scoring students (statistically significant at the 1 percent level; F-statistic = 900.64).[2] This instrumental variables regression produces a treatment-on-treated (TOT) estimate of the impact of double-dose on students induced into double-dose algebra by the introduction of the policy. Here, high-scoring students serve as a control group for low-scoring students, so that these estimates will be unbiased under the assumption that no changes other than double-dose algebra differentially affected low- and high-scoring students over this time period.

In addition, we further explore how the effects of the policy varied by student's academic skill. We reestimate the above regressions by interacting the instrument and endogenous regressor with three levels of math skill as measured by their incoming eighth grade test score: *very low* (below the twentieth percentile), *low* (between the twentieth and thirty-ninth percentiles), and *medium* (between the fortieth and forty-ninth percentiles).

II. Empirical Findings and Discussion

Before turning to the regression results, Figure 1 illustrates that tracking intensified under this policy. This figure graphs the mean eighth grade math percentile of each student's regular algebra class peers, averaged by the student's own eighth grade math percentile. Recall that the 2003–2004 cohorts were subject to the policy and the 2001–2002 cohorts were not. Figure 1 clearly shows that the increased tracking placed low-scoring students in algebra with lower-skilled peers and higher-scoring students with higher-skilled peers. Also, we note that no such sorting by math ability occurred in the pre-policy cohorts.

The results for the regression-adjusted DID analysis are reported in Tables 1 and 2. All coefficients reported come from DID regressions in which double-dosing has been instrumented by eligibility; these are TOT estimates of the impact of double-dosing on those students who were actually double-dosed. We present two sets of results: panel A shows the overall effect and panel B reports on the policy effects by math ability. These latter results are of policy interest because they identify the type of student who benefited the most from this intensive math curriculum, and thus provide insight on how to better target educational practices toward students of different skill levels.

Table 1 shows the impact of the policy on freshmen academic experience. Columns 1 and 2 in panel A, show that the policy increased the number of freshman math courses taken by about one, as would be expected from the double-dose strategy. Columns 3–6 highlight the policy's impact on channels other than increasing instructional time in math, namely the change in peers. Columns 3 and 4 imply that the policy lowered the mean peer skill of the average double-dosed student by 13 percentiles. Columns 5 and 6 suggest that double-dosed students were in more homogeneous classrooms than their non-double-dosed peers. The bulk of this effect is coming from students who are above the twentieth percentile of math skills (panel B, columns 11 and 12. In short, double-dosing increased instructional

[2] The first-stage results are provided in the online Appendix.

Table 1—Difference-in-Differences Regressions, The Effect of Double-Dose Algebra on Freshman Coursework and Peer Composition

	Math courses		Mean peer skill		SD of mean peer	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Overall						
Double-dosed	0.946***	0.940***	−13.001***	−13.306***	−1.230**	−1.197**
	(0.012)	(0.012)	(1.056)	(1.033)	(0.523)	(0.519)
μ (below $\times$ before)	0.96	0.96	41.59	41.59	18.02	18.02
	(7)	(8)	(9)	(10)	(11)	(12)
Panel B. By math ability						
Double-dosed $\times$ ($0 \leq$ math < 20)	0.830***	0.824***	−10.007***	−10.086***	−0.246	−0.237
	(0.024)	(0.024)	(1.173)	(1.142)	(0.509)	(0.509)
Double-dosed $\times$ ($20 \leq$ math < 40)	0.827***	0.821***	−11.448***	−11.796***	−0.978**	−0.942**
	(0.022)	(0.023)	(1.060)	(1.031)	(0.456)	(0.453)
Double-dosed $\times$ ($40 \leq$ math < 50)	0.840***	0.837***	−11.952***	−12.293***	−1.813***	−1.779***
	(0.025)	(0.025)	(1.118)	(1.121)	(0.513)	(0.509)
Observations	60,497	60,497	60,497	60,497	60,497	60,497
Controls:						
Student demographics		Yes		Yes		Yes
Neighborhood characteristics		Yes		Yes		Yes
High school fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses. Each column in each panel represents the instrumental variables difference-in-differences regressions of the listed outcome on a double-dosed indicator, where the first-stage regression is described in the text. Odd-numbered columns include only cohort and high school fixed effects, while even-numbered columns add controls for gender, race, free and reduced price lunch status, high school start age, cohort, eighth grade reading score, and census block poverty and socioeconomic measures. Also listed is the mean value of each outcome for students below the eligibility threshold in pre-policy years.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

time and homogeneity of classrooms but lowered the math skill of the average peer to which students were exposed.

We present the net impact of these changes on academic performance by focusing on two primary sets of outcomes: *short-* and *longer-term* impacts. First, in Table 2, we explore whether double-dose immediately helped students perform better on their freshman coursework. To measure the short-term gains, we construct two ninth grade measures of grade received in algebra and whether a student received a C or better in their algebra course. Second, we further investigate whether double-dose improves long-run outcomes. We analyze the math portion of the PLAN exam, which all CPS students take in September in the tenth grade; and construct math coursework performance indicators in later years and on time graduation.

As shown in Table 2, this policy raised the average double-dosed students' freshmen algebra GPA by 0.22 grade points (column 1) from a prior mean of 1.27. Double-dosing also increased the proportion of students earning a C or better in freshman algebra by 6.7 percentage points (column 2), a 25 percent improvement from a base mean of 27 percent. Next, columns 3 and 4 analyze the potential outcomes beyond freshman year. Remarkably, we observe positive effects of this policy on math coursework even two years after treatment. Double-dosed students were 3.8 percentage points more likely to receive at least a C in geometry by their second year and 3.2 percentage points more likely to get a C or higher in trigonometry by their third year of high school.

Though coursework and grades matter for students, academic trajectories, the subjective

Table 2— Difference-in-Differences Regressions, The Effect of Double-Dose Algebra on Math Coursework, Test Scores, and Attainment

	Algebra GPA (1)	Algebra, C or better (2)	Geometry, C or better (3)	Trigonometry, C or better (4)	Math test, z-score (5)	High school graduate (6)
Panel A. Overall						
Double-dosed	0.220***	0.067***	0.038*	0.032**	0.081***	0.017
	(0.050)	(0.019)	(0.021)	(0.014)	(0.024)	(0.014)
μ (below $\times$ before)	1.27	0.27	0.37	0.19	−0.34	0.43
	(7)	(8)	(9)	(10)	(11)	(12)
Panel B. By math ability						
Double-dosed $\times$ ($0 \leq$ math < 20)	0.075	0.042*	0.007	0.028	0.036	0.003
	(0.061)	(0.023)	(0.023)	(0.018)	(0.037)	(0.016)
Double-dosed $\times$ ($20 \leq$ math < 40)	0.222***	0.064***	0.042**	0.026**	0.096***	0.013
	(0.048)	(0.018)	(0.020)	(0.013)	(0.025)	(0.013)
Double-dosed $\times$ ($40 \leq$ math < 50)	0.211***	0.067***	0.038*	0.038**	0.058**	0.032*
	(0.048)	(0.019)	(0.020)	(0.017)	(0.030)	(0.019)
Observations	60,497	60,497	60,497	60,497	42,114	60,497
Controls:						
Student demographics	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes	Yes	Yes	Yes
High school fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Heteroskedasticity robust standard errors clustered by initial high school are in parentheses. Each column in each panel represents the instrumental variables difference-in-differences regressions of the listed outcome on a double-dosed indicator, where the first-stage regression is described in the text. All columns include cohort and high school fixed effects, as well as controls for gender, race, free and reduced price lunch status, high school start age, cohort, eighth grade reading score, and census block poverty and socioeconomic measures. Also listed is the mean value of each outcome for students below the eligibility threshold in pre-policy years. In columns 3 and 4, students who do not take the course are assigned zero as the outcome value.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

nature of course grading motivates us to turn to standardized measures of achievement and attainment. Columns 5 and 6 show that double-dosing improved math scores by a highly significant 0.08 standard deviations and improved graduation by a positive but statistically insignificant 1.7 percentage points.

Our final analysis looks at the heterogeneity effects of the policy by student's incoming math ability. The estimates in panel B of Table 2 show a clear pattern. Columns 7–12 show that the bulk of the positive impacts reported in panel A came through the policy's effect on low- and medium-skilled students and not on very low-skilled students. Most importantly, we also see larger and marginally significant positive impacts on high school graduation for students with medium math skills. Double-dosed students whose math ability were at or above the fortieth percentile experienced a 3.2 percentage point increase in graduation, a 6 percent improvement from a base rate of 54 percent for such students.

III. Concluding Remarks

In this article we present new evidence on the effects of ability-tracking, with one important distinction from many of the previous tracking studies. Despite the fact that double-dose students were channeled to "lower-track" algebra classes, these students also received twice the instructional time, more challenging coursework, and improved pedagogy than they would have received under traditional tracking. Specifically, the second math course (algebra with support), the double-dose teachers focused on building math skills that students lacked, used various instructional activities, such as working in a small group (cooperative groups), asking probing and open-ended questions. Hence, these specific aspects of the policy may have mitigated the potential negative effects of being placed with lower skilled peers. In contrast, under traditional tracking students are placed in more homogenous classes, but classes are often characterized as having low-level content and a poor instructional environment (Oakes 1985).

There are three main takeaways from the analysis here. First, we observe that this curricular intervention, which only occurred for one year, had lasting positive effects on longer-term student outcomes. Second, our results also reveal that the lowest skilled students benefitted less from the intervention than did somewhat higher skilled students, perhaps because the focus on building high-level analytical skills required a certain baseline math ability. Lastly, our results provide further evidence, consistent with Duflo, Dupas, and Kremer (2011), that the benefits of tracking and the resulting better-targeted pedagogy may outweigh the impact of being exposed to lower-skilled classmates.

REFERENCES

Abdulkadiroglu, Atila, Joshua D. Angrist, and Parag A. Pathak. 2011. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." National Bureau of Economic Research Working Paper 17264.

Betts, Julian R. 2011. "The Economics of Tracking in Education." In *Handbook of the Economics of Education*. Vol. 3, edited by Eric Hanushek, Stephen Machin, and Ludger Woessmann, 341–81. Amsterdam: North Holland.

Betts, Julian R., and Jamie L. Shkolnik. 2000. "The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools." *Economics of Education Review* 19 (1): 1–15.

Bui, Sa A., Steven G. Craig, and Scott A. Imberman. 2011. "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Achievement." National Bureau of Economic Research Working Paper 17089.

Clark, Damon. 2010. "Selective Schools and Academic Achievement." *B. E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy* 10 (1): Article 9.

Cortes, Kalena, Joshua Goodman, and Takako Nomi. 2013a. "A Double Dose of Algebra." *Education Next* 13 (1): 70–6.

Cortes, Kalena, Joshua Goodman, and Takako Nomi. 2013b. "Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra." Harvard Kennedy School Faculty Research Working Paper 13–09.

Dobbie, Will, and Roland G. Fryer, Jr. 2011. "Exam High Schools and Academic Achievement: Evidence from New York City." National Bureau of Economic Research Working Paper 17286.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101 (5): 1739–74.

Figlio, David N., and Marianne E. Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics* 51 (3): 497–514.

Lucas, Adrienne M., and Isaac Mbiti. Forthcoming. "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya." *American Economic Journal: Applied Economics*.

Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven: Yale University Press.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review* 103 (4): 1289–324.

Wheelock, Anne. 1992. *Crossing the Tracks: How "Untracking" Can Save America's Schools*. New York: New Press.

This article has been cited by:

1. Kamilah Legette. 2020. A Social-Cognitive Perspective of the Consequences of Curricular Tracking on Youth Outcomes. *Educational Psychology Review* **32**:3, 885-900. [Crossref]
2. João Batista Araujo e Oliveira, Matheus Gomes, Thais Barcellos. 2020. A Covid-19 e a volta às aulas: ouvindo as evidências. *Ensaio: Avaliação e Políticas Públicas em Educação* **28**:108, 555-578. [Crossref]
3. Se Woong Lee, Xinyi Mao. 2020. Algebra by the Eighth Grade: The Association Between Early Study of Algebra I and Students' Academic Success. *International Journal of Science and Mathematics Education* **115**. . [Crossref]
4. Serena Canaan. 2020. The long-run effects of reducing early school tracking. *Journal of Public Economics* **187**, 104206. [Crossref]
5. Ümmühan Yeşil Dağlı. 2019. Effect of increased instructional time on student achievement. *Educational Review* **71**:4, 501-517. [Crossref]
6. Shun-ichiro Bessho, Haruko Noguchi, Akira Kawamura, Ryuichi Tanaka, Koichi Ushijima. 2019. Evaluating remedial education in elementary schools: Administrative data from a municipality in Japan. *Japan and the World Economy* **50**, 36-46. [Crossref]
7. Philipp Mandel, Bernd Süssmuth, Marco Sunder. 2019. Cumulative instructional time and student achievement. *Education Economics* **27**:1, 20-34. [Crossref]
8. Emily C. Bouck, Jiyoon Park, Mary K Bouck, Jim Alspaugh, Stacey Spitzley. 2019. Exploration of a middle school Tier 2 math lab on student performance. *Preventing School Failure: Alternative Education for Children and Youth* **63**:1, 89-95. [Crossref]
9. David C. Geary, Daniel B. Berch, Kathleen Mann Koepke. Introduction: Cognitive Foundations for Improving Mathematical Learning 1-36. [Crossref]
10. Tyler W. Watts, Greg J. Duncan, Douglas H. Clements, Julie Sarama. 2018. What Is the Long-Run Impact of Learning Mathematics During Preschool?. *Child Development* **89**:2, 539-555. [Crossref]
11. Mathias Huebener, Susanne Kuger, Jan Marcus. 2017. Increased instruction hours and the widening gap in student performance. *Labour Economics* **47**, 15-34. [Crossref]
12. Drew Bailey, Greg J. Duncan, Candice L. Odgers, Winnie Yu. 2017. Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. *Journal of Research on Educational Effectiveness* **10**:1, 7-39. [Crossref]
13. Christian Krekel. 2017. Can Raising Instructional Time Crowd Out Student Pro-Social Behaviour? Evidence from Germany. *SSRN Electronic Journal* . [Crossref]
14. Emanuel Tamir, Lea Shaked. 2016. What to Do with the Bounty? Organizational Patterns for the Implementation of Resources Allocated by the Courage to Change (Oz Letmura) Reform. *Leadership and Policy in Schools* **15**:4, 567-597. [Crossref]
15. Simon Calmar Andersen, Maria Knoth Humlum, Anne Brink Nandrup. 2016. Increasing instruction time in school does increase learning. *Proceedings of the National Academy of Sciences* **113**:27, 7481-7484. [Crossref]
16. Susanne Kuger. Curriculum and Learning Time in International School Achievement Studies 395-422. [Crossref]
17. Mathias Huebener, Susanne Kuger, Jan Marcus. 2016. Increased Instruction Hours and the Widening Gap in Student Performance. *SSRN Electronic Journal* . [Crossref]
18. Vincenzo Andrietti. 2016. The Causal Effects of an Intensified Curriculum on Cognitive Skills: Evidence from a Natural Experiment. *SSRN Electronic Journal* . [Crossref]