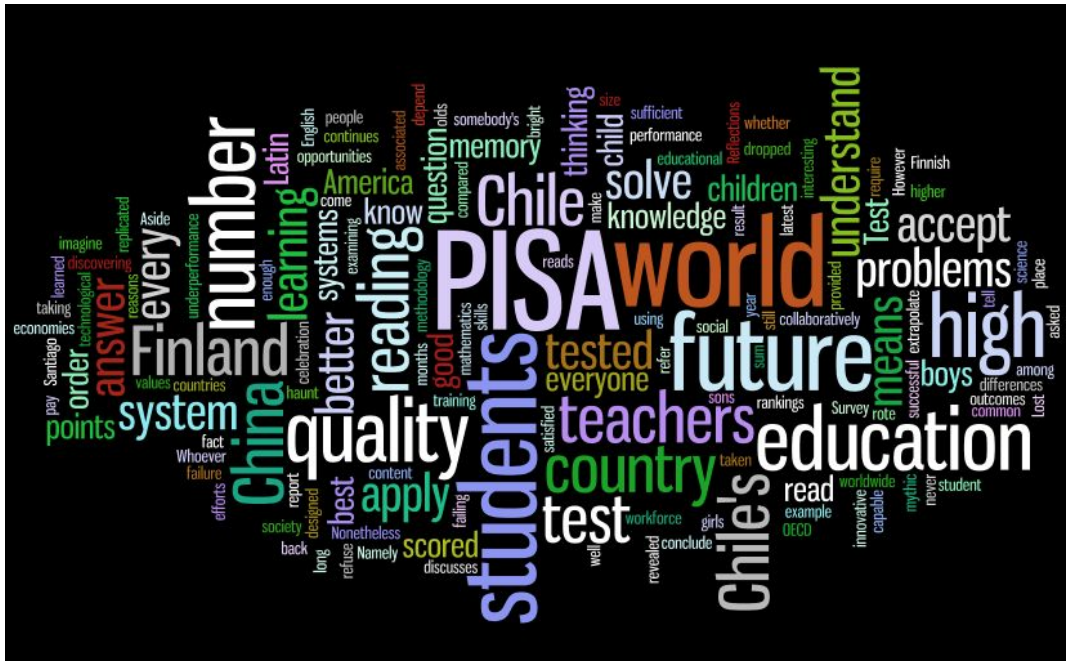
Revisiting the Relationship between International Assessment Outcomes and Educational Production: Evidence from a Longitudinal PISA-TIMSS Sample

Martin Carnoy[a]
Tatiana Khavenson[b]
Prashant Loyalka[a]
William Schmidt[c]
*Andrey Zakharov[b]

[a] Stanford University
[b] National Research University Higher School of Economics
[c] Michigan State University

In recent years, results from international student assessments (such as the PISA, the TIMSS) have gained in prominence



"The triennial results from the (PISA)…have become education's equivalent of the football World Cup"

– The Guardian

This increasing prominence of PISA and TIMSS results has placed pressure on policymakers to enact educational reforms

“It's normally parents, teachers and pupils who anxiously await exam results. But now [with the upcoming release of PISA results] it's the turn of education ministers to bite their nails.”

– The Guardian “The OECD’s Pisa Delivery Man”

“A NEW study from the [PISA] will land on the desks of policymakers around the world next month. It will make sobering reading for political leaders in many countries.”

– The Economist “Fixing Sweden’s Schools”

Along with their increasing prominence (especially in policy circles) the PISA and TIMSS results are also increasingly being used to recommend specific policies for reforming national education systems and schools

The Organization for Economic Cooperation and Development (OECD) has been especially active in using PISA results to recommend/promote national education and school-level reforms

OECD efforts are influential…

*"**PISA Day** is the start to an ongoing initiative that will look beyond the international rankings to learn how to improve assessments and education in the United States."*

- Bob Wise, President, Alliance for Excellent Education

(one of the hosts of PISA Day 2012)

ALLIANCE FOR EXCELLENT EDUCATION
Achieve ACT AMERICA ACHIEVES Asia Society Business Roundtable CCSSO CollegeBoard NATIONAL BOARD for Professional Teaching Standards
NCEE National Center On EDUCATION And The Economy

The OECD claims there will be strong repercussions if their arguments are not taken seriously…

“The generation born in the UK this year is likely to lose £4.5 trillion in economic output over their lifetime because UK schools aren't delivering what other countries' education systems show can be achieved…deficiencies in the UK's school systems amount to the equivalent of a permanent recession”

- Andreas Schleicher, OECD, 2012

Unfortunately, the OECD recommendations have been largely based on the policies and practices of those countries with the highest PISA scores…

Such recommendations are limited because of limitations inherent in the PISA data

In particular, the PISA data are beset by two problems that have long been recognized in the academic literature

First, the PISA scores (like the TIMSS scores) only reflect accumulated student achievement at one point in time

The scores thus not only represent the influence of school-related factors but also ***non-school-related*** factors. These factors are both observable and ***unobservable***

Disentangling the effects of school and non-school-related factors (observable and unobservable) can be a formidable task, when relying on data collected at a single point in time (Coleman, 1966)

Second, the survey strategy used by PISA makes it difficult to estimate the impact of classroom factors (e.g. teachers/OTL) on student achievement

- PISA randomly selects a small number (approximately 25) of 15 year-olds from each school in each sample and does not select intact classrooms

- PISA does not use a teacher questionnaire. PISA thus does not identify students with particular teachers, does not connect students to particular classrooms, to teaching practices, and to classroom curriculum implementation

The validity of the OECD's policy recommendations for raising student achievement can only be tested by data that allow us to resolve these two problems

The overall **goal** of our study

To test the robustness of the policy recommendations made by the OECD (using better data that do not have these problems)

Specifically, we seek to test the following OECD recommendations:

1) Hiring better (or more effective) teachers improves scores

2) Providing greater coverage and amount of time on subject matter ("opportunity to learn" or OTL) improves scores

3) Providing greater OTL helps disadvantaged students (e.g. lower social class) more than advantaged students (creates greater equality in learning outcomes)

(e.g. OECD, 2010, pp. 3-4; OECD, 2011; OECD, 2013a; Schleicher, 2014)

Approach: Data

- Unique data from a national sample of more than 4,000 Russian students that took the TIMSS test in 2011 and the PISA test (one year later) in 2012

- The data includes math scores on the same students at two points in time

- Because the TIMSS survey/test is classroom-based and because almost all Russian students remain with their fellow classroom students from the 8^{th} to the 9^{th} grade, we have information on teachers and intact classrooms in both 8^{th} (TIMSS) and 9^{th} grade (PISA)

Approach: Methodology

- We test the impact of teacher "quality" and OTL on PISA math scores using an educational production function approach - we control for baseline scores (ability), student background, teacher/ class/school factors

- Define teacher "quality" as:
  - Teacher pre-service education (math major or education major)
  - Teacher certification ("category") level

- Define OTL as:
  - exposure to applied mathematics concepts
  - exposure to work problems
  - exposure to formal mathematics concepts (algebra and geometry)

- We examine the relationship for the average student, by social class and by baseline achievement

Approach: Value-added model

- $A_{ijPISA2012} = a_1 A_{ijTIMSS} + \Sigma b'_{1k} X_{ijk} + b_2 AvgX_{ij} + \Sigma c'_{2k} TC_{jk} + c'_3 TExp_j + \Sigma c'_{4k} TEduc_{jk} + \Sigma d'_k TAct_{jk} + \Sigma f'_k OTL_{ijk} + \Sigma g'_k S_k + e'_{ij}$;

Results

Main Results

(for the average student)

VAM results – Teacher "quality"
(controlling for baseline achievement)

	II	III	IV	V
TIMSS Math Score 2011	0.53***	0.52***	0.53***	0.52***
Teacher Pre-Service Math Education	**-0.15****	**-0.15***	**-0.15***	**-0.15****
Teacher Pre-Service No Math	-0.17	-0.17	**-0.21***	-0.20
Years Teaching Math	0.01	0.00	0.00	0.00
Years Teaching Math Squared	-0.00	-0.00	-0.00	-0.00
Teacher "Highest" Category	0.00	0.00	-0.02	-0.03
Teacher "First" Category	0.04	0.04	0.03	0.02
Teacher "No Category"	-0.17	-0.22	-0.21	-0.23
Exposure Applied Math		-0.07***		
Exposure Word Problems			0.01	
Exposure Formal Math				0.09***
Constant	1.68**	1.17	1.22	1.11
Adjusted R^2	0.442	0.437	0.431	0.441

Controls: *age, gender, books in home, mom's education, teacher workload, school type*

*SEs adjusted for clustering at the school level, *** p<0.01, ** p<0.05, * p<0.1*

VAM results – Teacher "quality"
(controlling for baseline achievement)

	II	III	IV	V
TIMSS Math Score 2011	0.53***	0.52***	0.53***	0.52***
Teacher Pre-Service Math Education	-0.15**	-0.15*	-0.15*	-0.15**
Teacher Pre-Service No Math	-0.17	-0.17	-0.21*	-0.20
Years Teaching Math	0.01	0.00	0.00	0.00
Years Teaching Math Squared	-0.00	-0.00	-0.00	-0.00
Teacher "Highest" Category	0.00	0.00	-0.02	-0.03
Teacher "First" Category	0.04	0.04	0.03	0.02
Teacher "No Category"	-0.17	-0.22	-0.21	-0.23
Exposure Applied Math		**-0.07*** **		
Exposure Word Problems			0.01	
Exposure Formal Math				**0.09*** **
Constant	1.68**	1.17	1.22	1.11
Adjusted R^2	0.442	0.437	0.431	0.441

Controls: *age, gender, books in home, mom's education, teacher workload, school type*

*SEs adjusted for clustering at the school level, *** p<0.01, ** p<0.05, * p<0.1*

Subgroup Results

(by social class and achievement level)

VAM Results – teacher quality (by social class)

	Low			Mid			High		
	1	2	3	4	5	6	7	8	9
TIMSS Math Score 2011	0.44***	0.46***	0.44***	0.57***	0.57***	0.57***	0.59***	0.59***	0.59***
Teacher Pre-Service Math Education	-0.10	-0.09	-0.11	**-0.16***	**-0.17***	**-0.16***	**-0.16****	**-0.15****	**-0.16****
Teacher Pre-Service No Math	-0.19	-0.18	-0.19	-0.22	**-0.22***	-0.21	**-0.23***	**-0.22***	**-0.23***
Years Teaching Math	-0.00	-0.00	-0.00	0.01	0.01	0.01	-0.00	-0.00	-0.00
Years Teaching Math Squared	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Teacher "Highest" Category	-0.10	-0.12	-0.15	-0.03	-0.04	-0.05	0.15	0.14	0.12
Teacher "First" Category	-0.00	-0.01	-0.03	-0.04	-0.05	-0.05	0.16	0.16	0.14
Teacher "No Category"	-0.31	-0.31	-0.32	-0.26	-0.22	-0.26	-0.05	-0.05	-0.07
Exposure Applied Math	-0.11***			-0.07**			-0.05**		
Exposure Word Problems		-0.03			0.04			0.02	
Exposure Formal Math			0.10***			0.11***			0.05
Constant	1.03	1.10	1.20	2.28*	2.11	1.77	0.91	1.01	0.88
Adjusted R^2	0.310	0.297	0.312	0.447	0.441	0.453	0.491	0.489	0.492

VAM Results – OTL (by social class)

	Low			Mid			High		
	1	2	3	4	5	6	7	8	9
TIMSS Math Score 2011	0.44***	0.46***	0.44***	0.57***	0.57***	0.57***	0.59***	0.59***	0.59***
Teacher Pre-Service Math Education	-0.10	-0.09	-0.11	-0.16*	-0.17*	-0.16*	-0.16**	-0.15**	-0.16**
Teacher Pre-Service No Math	-0.19	-0.18	-0.19	-0.22	-0.22*	-0.21	-0.23*	-0.22*	-0.23*
Years Teaching Math	-0.00	-0.00	-0.00	0.01	0.01	0.01	-0.00	-0.00	-0.00
Years Teaching Math Squared	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Teacher "Highest" Category	-0.10	-0.12	-0.15	-0.03	-0.04	-0.05	0.15	0.14	0.12
Teacher "First" Category	-0.00	-0.01	-0.03	-0.04	-0.05	-0.05	0.16	0.16	0.14
Teacher "No Category"	-0.31	-0.31	-0.32	-0.26	-0.22	-0.26	-0.05	-0.05	-0.07
Exposure Applied Math	**-0.11*** **			**-0.07** **			**-0.05** **		
Exposure Word Problems		-0.03			0.04			0.02	
Exposure Formal Math			**0.10*** **			**0.11*** **			0.05
Constant	1.03	1.10	1.20	2.28*	2.11	1.77	0.91	1.01	0.88
Adjusted R^2	0.310	0.297	0.312	0.447	0.441	0.453	0.491	0.489	0.492

VAM Results – Teacher quality & OTL: students with different levels of baseline achievement

	TIMSS Benchmark 1-2			TIMSS Benchmark 3		
	4	5	6	7	8	9
TIMSS Math Score 2011	0.35***	0.35***	0.34***	0.31***	0.31***	0.31***
Teacher Pre-Service Math Education	-0.18	-0.20	-0.18	-0.13	-0.13	-0.15
Teacher Pre-Service No Math	-0.39	-0.41	-0.37	-0.14	-0.14	-0.17
Years Teaching Math	-0.02	-0.02	-0.02	0.00	0.00	0.00
Years Teaching Math Squared	0.00	0.00	0.00	0.00	0.00	0.00
Teacher "Highest" Category	-0.09	-0.08	-0.10	0.01	0.00	-0.03
Teacher "First" Category	-0.13	-0.13	-0.13	-0.00	-0.00	-0.04
Teacher "No Category"	**-0.47***	**-0.42***	**-0.48***	-0.38	-0.38	-0.38
Exposure Applied Math	-0.02			**-0.06***		
Exposure Word Problems		0.00			-0.03	
Exposure Formal Math			0.05			**0.10*****
Constant	0.97	1.05	0.96	2.87**	2.86**	2.64**
Adjusted R^2	0.179	0.182	0.184	0.122	0.116	0.136

VAM Results – Teacher quality & OTL: students with different levels of baseline achievement

	TIMSS Benchmark 4			TIMSS Benchmark 5		
	4	5	6	7	8	9
TIMSS Math Score 2011	0.31***	0.32***	0.32***	0.37***	0.38***	0.37***
Teacher Pre-Service Math Education	**-0.14***	**-0.13***	**-0.14***	-0.11	-0.09	-0.09
Teacher Pre-Service No Math	**-0.22***	-0.21	**-0.22***	-0.15	-0.09	-0.13
Years Teaching Math	0.01	0.01	0.01	0.01	0.01	0.01
Years Teaching Math Squared	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Teacher "Highest" Category	0.01	-0.01	-0.01	0.21	0.14	0.15
Teacher "First" Category	0.16	0.16	0.15	0.19	0.12	0.11
Teacher "No Category"	0.04	0.05	0.04	-0.04	-0.05	-0.07
Exposure Applied Math	**-0.07*****			**-0.11*****		
Exposure Word Problems		**0.05****			0.04	
Exposure Formal Math			**0.10****			**0.18*****
Constant	0.72	0.83	0.78	1.08	1.32	0.96
Adjusted R^2	0.147	0.141	0.149	0.328	0.313	0.331

Summary of main findings

First, students that have teachers with university degrees from mathematics programs (as opposed to degrees from education programs) score higher on the PISA

→ *This coincides with OECD recommendations and pre-existing, rigorous evidence from the academic literature*

Second, there are mixed results from increases in OTL

- Exposure to formal math concepts is positively related to PISA math scores
- Exposure to "applied mathematics concepts" or to "word problems" is NOT positively (or is even negatively) related to PISA math scores

→ *The small "effects" are at variance with the OECD and academic literature*

Third, neither university degree from mathematics programs nor OTL (exposure to formal math) help disadvantaged students…teacher quality helps high social class kids (or higher achieving kids) more

→ *The "effects" are the opposite of OECD claims*

Policy Implications

- It is not a good idea to use international test results, such as the PISA findings, to make sweeping generalizations about what works in education

- International assessments are useful in identifying broad trends, but they are no substitute for careful causal analyses carried out in particular national and social contexts

Thank you!