Effets de la réforme scolaire au Québec sur les compétences mathématiques des enfants du primaire et du secondaire.

Catherine Haeck[†], Pierre Lefebvre[††] and Philip Merrigan[††]

[†]Katholieke Universiteit Leuven et [††]Université du Québec à Montréal

CIQSS, novembre 2011

Introduction

General question: How does schooling matters?

- A consensus seems to have emerged from international surveys suggesting that a sizeable proportion of young people around age 15 in many countries do not appear to possess all of the skills required to meet the challenges of today's knowledge societies.
- Empirical research has shown that measures of schooling attainment alone may not be sufficient to capture the extent to which human capital triggers economic growth and impacts individual labour market outcomes (e.g. Currie and Thomas, 2001; Hanushek and Woessmann, 2008)
- Research shows that concrete measures of academic achievement and cognitive skills, along with educational attainment, are strongly correlated with labour market outcomes, such as earnings and unemployment.
- A number of studies have documented the specific importance of mathematical abilities in adulthood socioeconomic success (e.g. Murnane et al., 1995; Rose and Betts, 2004; Ingram and Neumann, 2006).

Introduction

Specific question: Can a competency-based approach increase the chances of success for all?

- The reform aimed at making schools more responsive to the changing needs of children in order to improve the chances of success for all.
- It moved teaching away from the traditional/academic approaches of memorization, drills and activity books, to a much more comprehensive approach focused on learning in a contextual setting in which children are expected to find the answers for themselves.
- We estimate the impact of Québec's ambitious and universal school reform implemented in the early 2000's on children's mathematical ability throughout primary (K-6) and secondary (7-11) school.

Agenda

1. Historical Context and Motivation
2. The Québec Reform
3. Methodology
4. Data set and Summary Statistics
5. Results and Discussion
6. Conclusion

International Context

Empirical research using international assessments of students in school highlights four main points:

1. there are significant international differences in overall tests scores, even among high-income countries.
2. there are important disparities in the results between students within the same country.
3. countries performing strongly generally display the smallest disparities in the results (Knighton, Brochu, and Gluszynski 2010; Gonzales et al., 2008).
4. social background is a strong determinant of student achievement in a number of countries (Fuchs and Woessmann, 2007; Bussière, Catwright and Knighton, 2004; Fleischman et al. 2010).

Improving the performance of low-skill students can help reduce the overall disparity in scores between well-off and deprived students, and increase the overall performance of the country.

Background Context and Motivation

In the United States, a number of reforms were implemented

1. Comprehensive school reform (CSR)
   - → The potential of CSR to improve children's performance remains unclear, with most studies showing only modest effects - or sometimes no effect - on student achievement (Vernez et al. 2006; Orland, Hoffman and Vaugh, 2010; Borman et al., 2003).
2. Charter schools
   - → Findings from recent studies on charter schools, (reviewed in Gleason et al., 2010) based on non-experimental methods, have also been mixed.

Background Context and Motivation

Research on CSR models and Charter Schools are limited in a number of ways.

1. A majority of studies assume that grant-receiving schools implement the model, yet research shows failure in implementing the full model (Vernez et al. 2004) over a number of years (Orland et al., 2010).
2. A majority of schools involved in these reforms exhibit higher than average poverty rates, such that results may not be transferable to lower poverty rate settings.
3. The variety of reforms implemented, the choice of students treated, the different possible financing mechanisms and the various geographical locations create considerable heterogeneity that is empirically difficult to address in order to provide convincing evidence.

Historical context of the reform

At the time of the reform, Québec was among the top performing countries in international assessments...

but it was still subject to severe criticism at home due to its alarmingly large high school dropout rate, especially among male students.

$\longrightarrow$ As of 2000, a comprehensive school reform impacting both primary and secondary schools was deployed all across the province of Québec.

Characteristics of this new approach

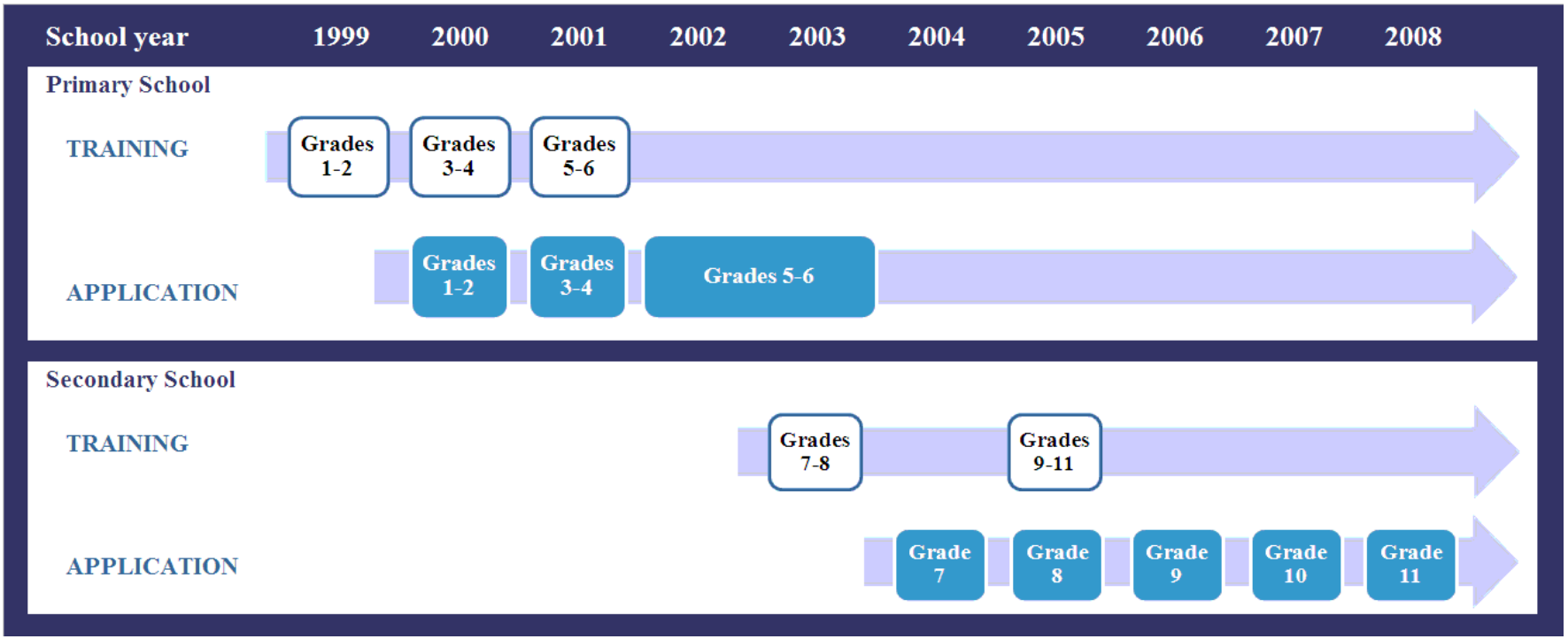
- focused on cross-curricular competencies and broad areas of learning (as opposed to specific knowledge)
- teaching and learning environment centered around the students to enable students to "find answers to questions arising out of everyday experience, to develop a personal and social value system, and to adopt responsible and increasingly autonomous behaviours" (MELS, 2005).
- students were expected to be more actively involved in their own learning and take responsibility for it. "Instead of passively listening to teachers, students will take in active, hands-on learning. They will spend more time working on projects, doing research and solving problems based on their areas of interest and their concerns. They will more often take part in workshops or team learning to develop a broad range of competencies." (MELS, 1999).

Characteristics of this new approach

Active competencies such as problem solving, strong communication skills, use of creativity, cooperation with others and teaching strategies based on the active participation of students were **central** to the reform...

while more **passive learning approaches** such as memorization, drill and traditional lectures in which teachers provide the content to be learned appears to have been put **aside**.

Reform schedule and implementation



International comparability

This approach was supported by a number of countries
$\longrightarrow$ e.g. reform-oriented teaching approach (United States) supported by leading organizations such as

- the National Council of Teachers of Mathematics
- the National Research Council
- the American Association for the Advancement of Science

Yet it remains unclear whether the traditional approach is preferable or not to the contextual approach focused on the development of competencies.

A small US study...

Crawford and Snider (2000) study two curricula (2 years, grade 4 to 5):

1. explicitly teaches mathematical concepts and enforces mastery through drill and repetition
2. more implicit approach in which students have to learn and discover concepts

Empirical approach: experiment with 46 students allocated randomly 23/23.

Finding

The more explicit approach was more successful in producing math knowledge

Another US study...

Le et al. (2006) study reform-oriented teaching on math achievements (3 years, grade 3 to 9).
Empirical approach: Multivariate analysis on about 1,600 to 3,500 students in 3 districts.

Limits of their study:

- No random assignment: students experienced a mix of approaches over the period.
- Teachers self-selected into the implementation of the approach.
- Teaching approach was self-reported and authors find conflicting evidence.
- Teachers admitted being influenced by the testing environment.

Finding

Results are at best weakly positive.

Methodological advantages of the reform

The Québec experiment/reform provides some advantages for the purpose of evaluation and cuts across some of the methodological difficulties mentioned above.

1. Québec's Department of Education implemented the reform and all schools (public and private) were forced to apply the new education program.
2. at the same time teaching in the rest of Canada continued to be delivered in the same way.
3. the reform's curriculum content was supported by a number of countries.
4. the data set used provides an observation period that is longer than typically seen in the literature.

Empirical Strategy

2 groups, 2 periods
$G_i \in \{0, 1\}$: group 0 is the RofC and group 1 is Québec
$T_i \in \{0, 1\}$: period 0 is prior to the school reform and period 1 is after

Y is the outcome measured

- Y^I_{gt} is the outcome if treated
- Y^N_{gt} is the outcome if *not* treated

We observe Y^I_{11}, Y^N_{10}, Y^N_{01}, Y^N_{00}

Empirical Strategy

The obvious candidate:

- the difference-in-differences (DID) estimator...

$\longrightarrow$ but DID has raised a number of concerns (e.g. Bertrand, Duflo, and Mullainathan, 2004; Donald and Lang, 2007; and Besley and Case, 2000).

As a result, we also implement

- the changes-in-changes (CIC) model (Athey and Imbens, 2006)
- the matching difference-in-differences (MDID) estimator (Heckman, Ishimura and Todd, 1997)

The DID approach

We first implement the following difference-in-differences estimator:

$$\tau^{DID} \equiv E[Y_{11}^I] - E[Y_{11}^N] = (E[Y_{11}] - E[Y_{10}]) - (E[Y_{01}] - E[Y_{00}]).$$

Assumptions:

- outcomes are additive in time period, group and unobservable characteristics of the individual (linearity)
- the treatment effect is constant across individuals or the effect differs across individuals but the distribution of outcomes without treatment is common across groups

The CIC approach

We implement the following changes-in-changes estimator:

$$\tau^{CIC} \equiv E\left[Y_{11}^{I}\right] - E\left[Y_{11}^{N}\right] = E\left[Y_{11}^{I}\right] - E\left[F_{Y,01}^{-1}\left(F_{Y,00}\left(Y_{10}\right)\right)\right].$$

Assumption:

- the underlying production functions for treated individuals and non-treated individuals, mapping the relationship between the outcomes and the unobservables at a given point in time, do not vary across groups

This model relaxes some of the assumptions of the standard DID:

- nonparametric identification, estimation, and inference for the ATE
- the time and treatment effect may vary across individuals
- estimates the entire counterfactual distribution of effects of not receiving the treatment for the treatment group
- it accommodates the possibility – but does not assume – that the treatment group adopted the policy because it expected greater benefits than in the control group

The MDID approach

In our approach to CIC, we control for X through a linear specification. To address the possibility of non linearity of response with respect to X, we implement the following matching difference-in-differences estimator:

$$\tau^{MDID} = \sum_{i \in S_{11}} \left\{ \left[y_{it_1} - \sum_{j \in S_{10}} \tilde{w}_{ijt_0} y_{jt_0} \right] - \left[\sum_{j \in S_{01}} \tilde{w}_{ijt_1} y_{jt_1} - \sum_{j \in S_{00}} \tilde{w}_{ijt_1} y_{jt_1} \right] \right\} w_i .$$

In contrast to standard DID, MDID allows the possibility of

- non linearity of response with respect to X
- selection into treatment

*** We implement kernel matching, local linear regression matching and nearest neighbor matching. Bootstrap standard errors are calculated for local linear regression and kernel matching to account for the underlying matching procedure (not consistent for nearest neighbor).

The Data set

We use Statistics Canada's National Longitudinal Survey of Children and Youth (NLSCY) which provides three cohorts of children of primary and secondary school age:

1. students in grades 1 to 6 in academic year 1994-95 up to grades 9 to 12 in academic year 2008-09
2. students in grades 1 to 4 in academic year 2006-07
3. students in grades 1 and 2 in academic year 2008-09

The NLSCY provides one measure of cognitive development for school age children: the CAT/2 mathematics test.

Characteristics of the test

- developed after careful consideration of the differences across the main school curricula across Canada
- designed to measure basic skills in mathematics (addition, subtraction, multiplication, division on integers, etc.)
- administered to students enrolled in grades 2 to 10, aged 7 to 15 years old
- difficulty of the test varies with the school grade of the child
- the standardized scores are designed to numerically represent the relative level of mathematics a child has attained

Points of caution...

1. The response rates for waves 1 to 3 were uncharacteristically low: 51% in wave 1, 74% in wave 2, and 54% in wave 3.
2. The difficulty level of the test for comparable students is different in wave 1 (compared to all other waves).
3. The difficulty level of the test for grades 9 and 10 is only comparable from wave 5 onwards.

School grades observed pre and post reform by cohort

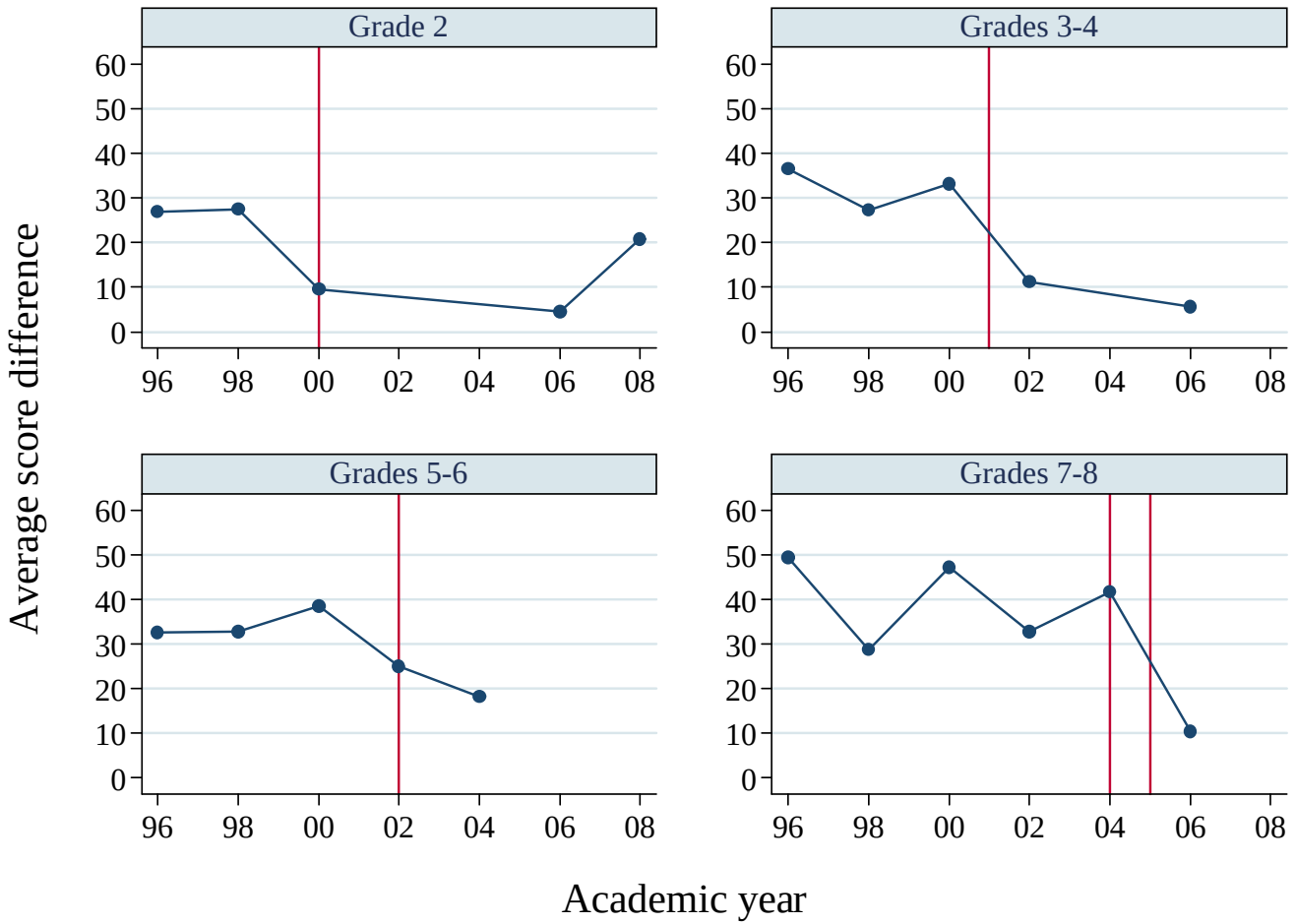
	Academic Year							Years in
Cohort	1996	1998	2000	2002	2004	2006	2008	reform
1	5 - 6	7 - 8	9 - 10					0
2	3 - 4	5 - 6	7 - 8	9 - 10				0
3	2	3 - 4	5 - 6	7 - 8	9 - 10			0
4		2	3 - 4	5 - 6	7 8	9 10		0, 1, 2, 4, 6
5			2	3 - 4	5 - 6	7 - 8	9 - 10	1, 3, 5, 7, 9
6				na	na	3 - 4	na	3, 4
7						2	na	2
8							2	2

Summary statistics

	Before		After	
	Mean	Std. dev.	Mean	Std. dev.
Student characteristics				
male	0.51	0.50	0.50	0.50
school grade				
2	0.13	0.33	0.18	0.38
3 and 4	0.16	0.37	0.26	0.44
5 and 6	0.15	0.36	0.30	0.46
7 and 8	0.30	0.46	0.14	0.35
9 and 10	0.27	0.44	0.12	0.32
ppvt (age 4-5)	100.34	14.65	99.87	15.12
math CAT/2				
grade 2	310.53	45.95	285.42	40.68
grades 3 and 4	367.12	54.25	359.22	51.18
grades 5 and 6	441.13	57.94	434.78	55.09
grades 7 and 8	502.84	71.05	487.44	68.94
grades 9 and 10	589.41	89.92	596.65	87.49
Family characteristics				
family structure				
one parent	0.19	0.39	0.18	0.39
two parents	0.81	0.39	0.81	0.39
household income ('000s)	71.53	53.47	84.32	64.60
mother works (dummy)	0.83	0.37	0.85	0.36
maternal education				
less than secondary	0.12	0.33	0.10	0.30
secondary	0.26	0.44	0.23	0.42
some post-secondary	0.20	0.40	0.14	0.35
college or university	0.41	0.49	0.52	0.50
area of residence				
rural	0.13	0.33	0.13	0.34
urban, $\leqslant$30,000	0.21	0.40	0.16	0.37
urban, 30,000 to 99,999	0.09	0.28	0.10	0.30
urban, 100,000 to 499,999	0.18	0.38	0.17	0.38
urban, $\geqslant$500,000	0.40	0.49	0.44	0.50
Nbr. of obs.	10,268		19,537	

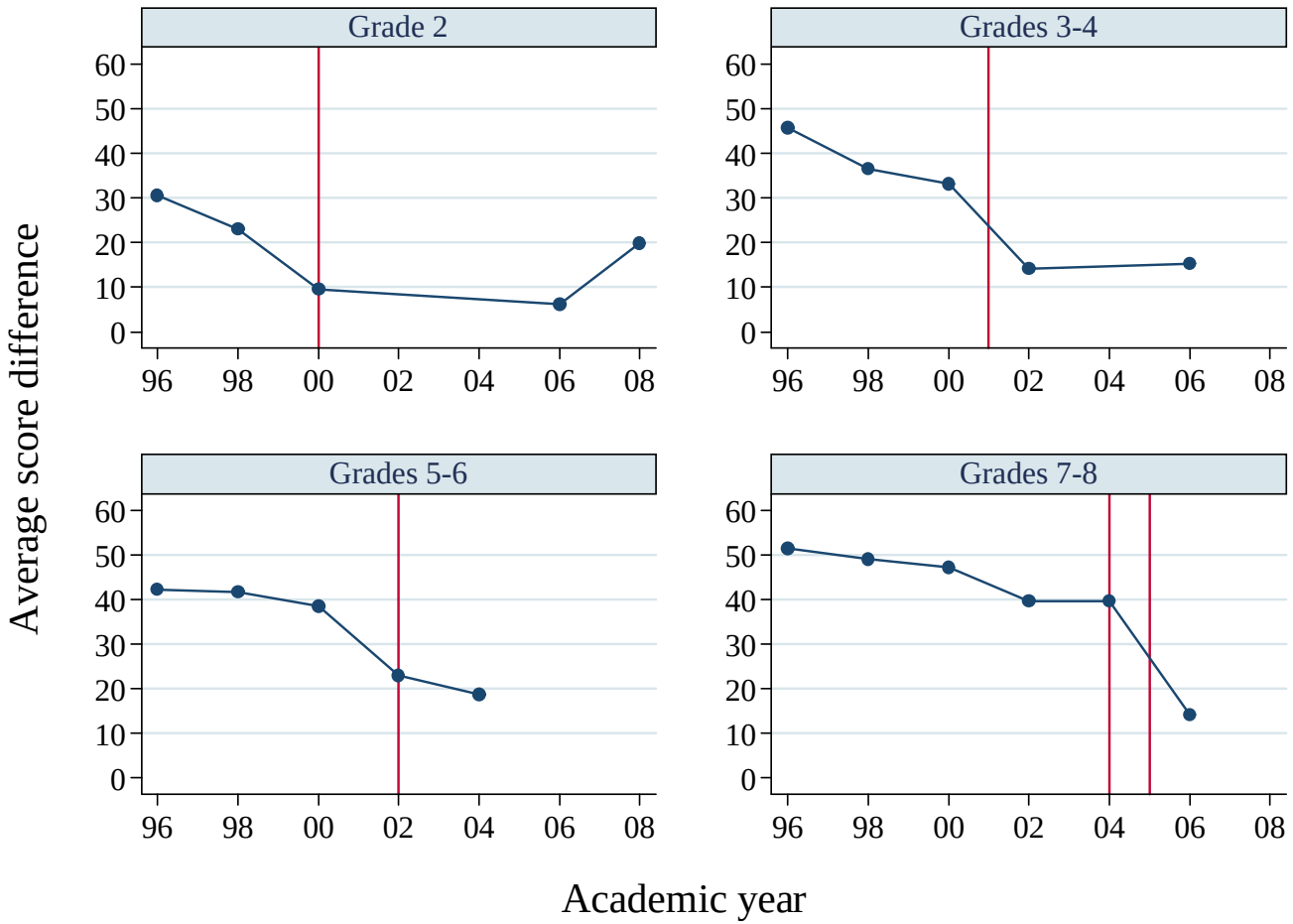
Trends in average score differences

No matching



Trends in average score differences

Matching



DID and MDID Estimates

	DID		DID with cov.		MDID llr		MDID kernel		MDID neighbor(5)	
	Mean	(Std.err.)	Mean	(Std.err.)	Mean	(Std.err.)	Mean	(Std.err.)	Mean	(Std.err.)
GRADE 2										
AcYear 1996, 2000*	-17.321	(8.383)	-21.182	(8.171)	-15.676	(3.492)	-14.129	(4.440)	-13.746	(4.811)
AcYear 1996, 2006	-22.285	(8.365)	-24.307	(8.093)	-22.389	(3.007)	-25.402	(3.577)	-15.624	(3.791)
AcYear 1996, 2008	-6.062	(8.548)	-11.067	(8.372)	-9.368	(4.301)	-10.152	(5.220)	-6.178	(6.537)
GRADES 3-4										
AcYear 2000, 2002*	-21.969	(5.025)	-19.636	(4.431)	-11.918	(2.461)	-7.077	(2.848)	-6.566	(2.971)
AcYear 2000, 2006	-27.604	(5.274)	-16.656	(4.870)	-11.053	(3.224)	-3.379	(3.185)	-3.876	(3.552)
GRADES 5-6										
AcYear 2000, 2002	-13.403	(6.230)	-13.428	(5.968)	-13.770	(3.134)	-9.186	(3.778)	-18.042	(4.064)
AcYear 2000, 2004*	-20.300	(6.118)	-20.126	(5.855)	-19.098	(3.022)	-20.805	(3.450)	-21.348	(3.835)
GRADES 7-8										
AcYear 2000, 2006*	-36.856	(7.814)	-33.428	(7.459)	-30.998	(4.468)	-31.967	(4.967)	-28.530	(5.836)
AcYear 2002, 2006*	-22.471	(7.522)	-23.079	(6.695)	-26.288	(4.451)	-32.465	(5.404)	-22.321	(6.121)
GRADES 9-10										
AcYear 2002, 2008*	-51.532	(12.413)	-45.121	(10.749)	-23.930	(6.239)	-34.609	(8.289)	-31.444	(7.386)
AcYear 2004, 2008*	-23.977	(11.466)	-28.584	(10.730)	-26.394	(6.186)	-29.172	(7.081)	-22.789	(6.850)

CIC Estimates - Grades 2 to 4

	Mean	(Std.err.)	25th Perc.	(Std.err.)	50th Perc.	(Std.err.)	75th Perc.	(Std.err.)	90th Perc.	(Std.err.)
GRADE 2										
Years 1996, 2000 (cohort 5)										
CIC disc ci	-17.039	(7.509)	-14.000	(8.967)	-10.000	(12.602)	-28.000	(12.438)	-13.000	(13.362)
CIC disc lower	-17.806	(7.534)	-14.000	(9.083)	-10.000	(12.653)	-28.000	(12.929)	-13.000	(14.122)
CIC disc upper	-16.476	(7.502)	-14.000	(8.918)	-10.000	(12.604)	-25.000	(12.342)	-13.000	(13.360)
Years 1996, 2006 (cohort 7)										
CIC disc ci	-21.215	(7.998)	-8.000	(6.840)	-14.000	(14.843)	-37.000	(14.737)	-31.000	(13.249)
CIC disc lower	-22.020	(7.997)	-9.000	(6.960)	-14.000	(14.809)	-37.000	(14.705)	-39.000	(13.306)
CIC disc upper	-20.582	(7.992)	-8.000	(6.824)	-14.000	(14.891)	-37.000	(14.848)	-28.000	(13.553)
Years 1996, 2008 (cohort 8)										
CIC disc ci	-5.428	(7.918)	2.000	(7.310)	-1.000	(13.356)	-18.000	(13.480)	0.000	(18.143)
CIC disc lower	-6.233	(7.968)	-1.000	(7.490)	-1.000	(13.521)	-18.000	(13.513)	-5.000	(18.342)
CIC disc upper	-4.831	(7.888)	2.000	(7.270)	-1.000	(13.305)	-18.000	(13.634)	0.000	(18.493)
GRADES 3-4										
Years 2000, 2002 (cohort 5)										
CIC disc ci	-15.171	(4.184)	-16.000	(5.913)	-17.000	(4.770)	-11.000	(4.760)	-15.000	(8.098)
CIC disc lower	-15.592	(4.194)	-16.000	(6.044)	-18.000	(4.900)	-11.000	(4.905)	-15.000	(8.194)
CIC disc upper	-14.711	(4.178)	-16.000	(5.923)	-17.000	(4.822)	-11.000	(4.738)	-15.000	(8.083)
Years 2000, 2006 (cohort 6)										
CIC disc ci	-16.722	(5.008)	-21.000	(6.288)	-16.000	(7.254)	-9.000	(5.837)	-11.000	(7.869)
CIC disc lower	-17.215	(5.013)	-22.000	(6.303)	-16.000	(7.354)	-9.000	(5.984)	-11.000	(7.941)
CIC disc upper	-16.144	(5.008)	-21.000	(6.300)	-16.000	(7.196)	-9.000	(5.780)	-9.000	(7.872)

CIC Estimates - Grades 5 to 10

	Mean	(Std.err.)	25th Perc.	(Std.err.)	50th Perc.	(Std.err.)	75th Perc.	(Std.err.)	90th Perc.	(Std.err.)
GRADES 5-6										
Years 2000, 2002 (cohort 4)										
CIC disc ci	-9.467	(6.052)	-10.000	(5.615)	0.000	(7.799)	-17.000	(10.626)	-16.000	(11.439)
CIC disc lower	-10.025	(6.069)	-10.000	(5.652)	0.000	(7.888)	-17.000	(10.675)	-16.000	(11.341)
CIC disc upper	-8.840	(6.041)	-10.000	(5.649)	0.000	(7.776)	-17.000	(10.788)	-16.000	(11.550)
Years 2000, 2004 (cohort 5)										
CIC disc ci	-19.460	(6.225)	-13.000	(6.830)	-14.000	(9.030)	-35.000	(10.323)	-25.000	(11.801)
CIC disc lower	-20.092	(6.230)	-14.000	(6.977)	-14.000	(9.085)	-35.000	(10.452)	-25.000	(11.893)
CIC disc upper	-18.851	(6.218)	-13.000	(6.766)	-14.000	(9.020)	-32.000	(10.238)	-24.000	(11.768)
GRADES 7-8										
Years 2000, 2006 (cohort 5)										
CIC disc ci	-29.824	(7.363)	-28.000	(11.243)	-39.000	(10.277)	-22.000	(10.740)	-6.000	(14.652)
CIC disc lower	-30.284	(7.382)	-28.000	(11.252)	-39.000	(10.303)	-22.000	(10.792)	-19.000	(14.731)
CIC disc upper	-29.240	(7.344)	-24.000	(11.261)	-39.000	(10.299)	-21.000	(10.699)	-6.000	(14.690)
Years 2002, 2006 (cohort 5)										
CIC disc ci	-23.738	(7.344)	-25.000	(7.577)	-24.000	(8.754)	-12.000	(11.200)	-1.000	(12.784)
CIC disc lower	-24.351	(7.360)	-26.000	(7.547)	-26.000	(8.852)	-12.000	(11.201)	-5.000	(13.128)
CIC disc upper	-23.129	(7.333)	-25.000	(7.606)	-23.000	(8.750)	-12.000	(11.263)	-1.000	(12.682)
GRADES 9-10										
Years 2002 2008 (cohort 5)										
CIC disc ci	-43.508	(11.105)	-66.000	(23.374)	-46.000	(16.775)	-47.000	(15.852)	-24.000	(17.089)
CIC disc lower	-43.907	(11.109)	-66.000	(23.378)	-48.000	(16.876)	-47.000	(15.843)	-24.000	(17.107)
CIC disc upper	-43.127	(11.102)	-65.000	(23.456)	-46.000	(16.780)	-47.000	(15.863)	-24.000	(17.130)
Years 2004, 2008 (cohort 5)										
CIC disc ci	-26.870	(10.287)	-37.000	(14.137)	-24.000	(15.115)	-27.000	(14.215)	-16.000	(15.674)
CIC disc lower	-27.449	(10.277)	-37.000	(14.165)	-24.000	(15.151)	-27.000	(14.140)	-17.000	(15.556)
CIC disc upper	-26.369	(10.300)	-36.000	(14.182)	-24.000	(15.234)	-27.000	(14.403)	-16.000	(15.848)

Impact of exposure

Focusing on cohort 5, we find that the magnitude of the effect is **increasing with exposure** to the reform (except from grade 2 to grades 3-4).
Grade 2: the mean effect is 17.0 (37.0% of a std. dev.)
Grades 3-4: the mean effect is 15.2 (28.0% of a std. dev.)
Grades 5-6: the mean effect is 19.5 (33.7% of a std. dev.)
Grades 7-8: the mean effect is 23.7 to 29.8 (33.3% to 41.9% of a std. dev.)
Grades 9-10: the mean effect is 26.9 to 43.5 (29.9% to 48.4% of a std. dev.)
$\longrightarrow$ Estimated impact of treatment on cohort 4 and 5 students in grade 5-6 also support this idea (1 year versus 5 years, -9.5 n.s. versus -19.5 s.)

Age at first exposure

Comparing students spending 1 to 2 years in the reform in grades 5-6 (cohort 4) with those in grade 2 (cohorts 5 and 7) it appears that the reform had a **greater impact on younger** children since the estimated impacts are negative and significant for them, while it is not different from zero for older students in grades 5-6.
*** This finding needs to be interpreted with caution, as estimated effects on grade 2 students rely on observations with higher non response.

Short term versus long term effects

We find that grade 2 students, **8 years after** the implementation of the reform, no longer seem to experience a significant negative effect (the CIC estimator for academic years 2008 is small and not different from zero). Possible explanations:

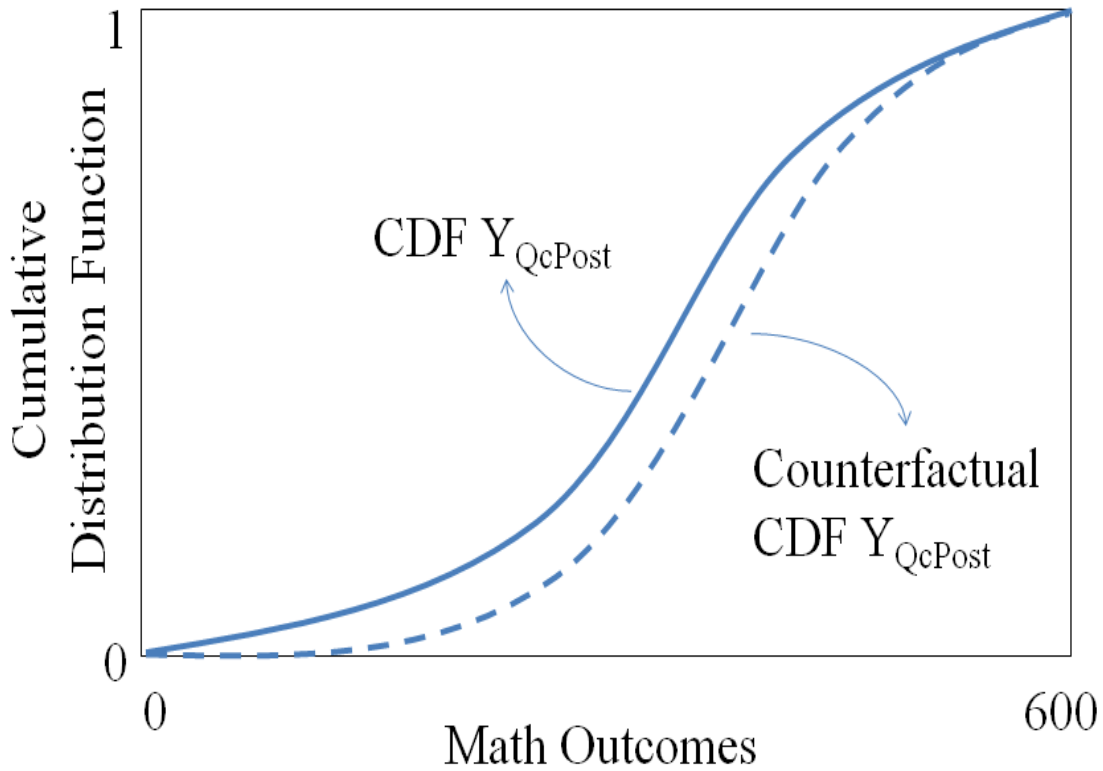
- took a fair number of years for teachers to develop the necessary skills to fully deploy all aspects of the reform.
- observing the decline in students' academic performance, teachers informally decided to reintroduce some of their pre-reform teaching approaches, and set aside in part or in totality the reform approach.

***This conclusion is derived from one set of grade 2 students at one point in time. Further research is needed to fully understand the long term effects of the reform.

Across the skill distribution

Results hold true for

- lower and middle performing students
- top performing students (but often not significant)



The entire distribution was impacted negatively.

PISA performance across provinces - Reading & Math

	2000		2003		2006		2009	
	Mean	(Std. err.)	Mean	(Std. err.).	Mean	(Std. err.)	Mean	(Std. err.)
Reading (global scores)								
Newfoundland	517	(2.8)	521	(4.9)	514	(5.4)	506	(6.2)
Prince Edward Island	517	(2.4)	495*	(4.4)	497*	(5.1)	486*	(5.5)
Nova Scotia	521	(2.3)	513	(4.4)	505*	(5.7)	516	(5.6)
New Brunswick	501	(1.8)	503	(4.3)	497	(5.0)	499	(5.5)
Québec	**536**	**(3.0)**	**525**	**(5.7)**	**522**	**(6.7)**	**522***	**(5.8)**
Ontario	533	(3.3)	530	(5.1)	534	(6.4)	531	(5.8
Manitoba	529	(3.5)	520	(5.0)	516	(5.7)	495*	(6.1)
Saskatchewan	529	(2.7)	512*	(5.6)	507*	(6.3)	504*	(5.9)
Alberta	550	(3.3)	543	(5.7)	535	(6.1)	533*	(6.7)
British Columbia	538	(2.9)	535	(4.5)	528	(7.1)	525	(6.5)
All of Canada	534	(1.6)	528	(4.1)	527	(5.1)	524	(5.2)
Mathematics (global scores)								
Newfoundland			517	(2.5)	507	(3.1)	503*	(3.4)
Prince Edward Island			500	(2.0)	501	(2.7)	487*	(3.0)
Nova Scotia			515	(2.2)	506	(2.8)	512	(3.0)
New Brunswick			512	(1.8)	506	(2.5)	504*	(3.0)
Québec			**537**	**(4.7)**	**540**	**(4.4)**	**543**	**(3.9)**
Ontario			530	(3.6)	526	(4.0)	526	(3.8)
Manitoba			528	(3.1)	521	(3.6)	501*	(4.1)
Saskatchewan			516	(3.9)	507	(3.7)	506	(3.8)
Alberta			549	(4.3)	530*	(4.0)	529*	(4.8)
British Columbia			538	(2.4)	523*	(4.7)	523*	(5.0)
All of Canada			532	(1.8)	527	(2.4)	527	(2.6)

PISA performance across provinces - Sciences

	2000		2003		2006		2009	
	Mean	(Std. err.)	Mean	(Std. err.).	Mean	(Std. err.)	Mean	(Std. err.)
Science (global scores)								
Newfoundland					526	(2.5)	518	(3.9)
Prince Edward Island					509	(2.7)	495*	(3.5)
Nova Scotia					520	(2.5)	523	(3.7)
New Brunswick					506	(2.3)	501	(3.5)
Québec					**531**	**(4.2)**	**524**	**(4.1)**
Ontario					537	(4.2)	531	(4.2)
Manitoba					523	(3.2)	506*	(4.8)
Saskatchewan					517	(3.6)	513	(4.5)
Alberta					550	(3.8)	545	(4.9)
British Columbia					539	(4.7)	535	(4.8)
All of Canada					534	(2.0)	529	(3.0)

Note that PISA results may be upward biased due to Québec's higher non-response rate in 2009 among students from less favorable socioeconomic environments.

TIMSS performance across provinces

	Mathematics Achievement Grade 4				Mathematics Achievement Grade 8			
Year	1995	1999	2003	2007	1995	1999	2003	2007
International	500	500	500	500	500	500	500	500
Québec	550	-	506	519	556	566	543	528
Ontario	489	-	511	512	501	517	521	517
Alberta	523	-	-	505	-	-	-	-
British Columbia	-	-	-	505	-	522	-	509
	Science Achievement Grade 4				Science Achievement Grade 8			
Year	1995	1999	2003	2007	1995	1999	2003	2007
International	500	500	500	500	500	500	500	500
Québec	529	-	500	517	510	540	531	507
Ontario	516	-	540	536	496	518	533	536
Alberta	555	-	-	543	-	-	-	-
British Columbia	-	-	-	537	-	542	-	526

Conclusion

Main findings:

- Negative impact across the entire distribution
- The longer the exposure the stronger the impact
- The younger at first exposure the more detrimental
- Long run effects may be nulle

Since the least performing students were impacted negatively, it appears that the reform failed to meet its main objective which was to raise the proportion of students who were successful in school.
Limits:

- Other outcomes may be important: reading, science, behaviour.
- Long term effects may be different... only one grade 2 cohort

*** Nonetheless, math achievement are important in predicting labour market outcomes.

Future research

- Impact on graduation rate (2010-2011 post reform) and high school drop out rates (in 5 years)
- Labour market effect on entrepreunership... to be observed in 5 to 10 years

MERCI

Titre du papier disponible en ligne:

All students left behind: An ambitious provincial school reform in Canada, but poor math achievements from grade 2 to 10.