

Felteau, Claude; Lefebvre, Pierre

## Working Paper

Can universal preschool education intensities counterbalance parental socioeconomic gradients? Repeated international evidence from Fourth graders skills achievement

Research Group on Human Capital - Working Paper Series, No. 23-01

## Provided in Cooperation with:

University of Quebec in Montreal, School of Management Sciences (ESG UQAM), Research Group on Human Capital (GRCH)

*Suggested Citation:* Felteau, Claude; Lefebvre, Pierre (2023) : Can universal preschool education intensities counterbalance parental socioeconomic gradients? Repeated international evidence from Fourth graders skills achievement, Research Group on Human Capital - Working Paper Series, No. 23-01, Université du Québec à Montréal, École des sciences de la gestion (ESG UQAM), Groupe de recherche sur le capital humain (GRCH), Montréal

This Version is available at:

<https://hdl.handle.net/10419/299316>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



## Research Group on Human Capital Working Paper Series

Can universal preschool education intensities counterbalance  
parental socioeconomic gradients?  
Repeated international evidence from Fourth graders skills achievement

Working Paper No. 23-01

Claude Felteau and Pierre Lefebvre

February 2023



Groupe de recherche sur le  
**CAPITAL HUMAIN**  
**ESG UOAM**

<https://grch.esg.uqam.ca/en/working-papers-series/>

# Can universal preschool education intensities counterbalance parental socioeconomic gradients? Repeated international evidence from Fourth graders skills achievement

Claude Felteau and Pierre Lefebvre  
Department of Economics and Research Group on Human Capital,  
School of Management, Université du Québec à Montréal (ESG UQAM)

February 2023

## Abstract

This study estimates the average multivalued treatment effects (ATET), of preschool attendance measured in years, on students' international *reading, math and science test z-scores in Grade 4*. The causal treatment effects come from multiple-years observational data on three levels of preschool duration before entering Grade 1. Among European countries that participated in five international education surveys, PIRLS (2006, 2011, 2016) and TIMSS (2015, 2019), those renowned for having adopted early childhood education (ECE) programs starting at a young age, growing in intensity and improving the number of qualified child-care providers were selected. In addition to four Nordic countries (Denmark, Finland, Norway, Sweden), France, two Belgium jurisdictions (French, Flemish), and two participating Canadian provinces, Ontario and Québec, were retained. The approach exploits the repeated surveys and cross-national comparative international z-scores tests. The data sets besides their test scores provide unique information from a parent questionnaire on their education and occupation levels, literacy and numeracy preschool activities, on child preschool educational childcare span in years and two program types (for some years; before and after age 3).

Four key findings can be identified from the data sets and estimations. First, there are large differences in the average scale score and percentiles deviation when converted into the z-score metric, for all categories of test scores across jurisdiction participants, and over time. Second, the estimates of the preschool treatment effects display rather heterogeneous impacts on z-scores with increasing significant and positive achievements over year surveys. Third, in general, preschool treatment effects are scattered in function of duration, programs types, and parental education. Four, results highlight stark gaps in scores related to parental education, socioeconomic statuses, and home learning resources for all year-samples. Evidence from a diversity of estimated gradients suggests established social inequalities in education achievement at ages 9-10 in Grade 4 could be difficult to reverse, even in cases where preschool education and care are implemented at a very young age in rich countries with very generous family policies.

**Code JEL code:** I2

**Key words:** Preschool education years intensities; fourth graders reading, math and science tests scores; multivalued treatment effects; PIRLS (2006, 2011, 2016); TIMSS (2015, 2019); Denmark, Finland, Norway, Sweden, Belgium French, Belgium Flemish, France; Canadian provinces of Québec and Ontario.

---

Contact: [Lefebvre.Pierre@uqam.ca](mailto:Lefebvre.Pierre@uqam.ca), Department of Economics, UQAM, CP8888, « Centre-ville », Montréal, Québec, Canada H3C 3P8

#We gratefully acknowledge the financial support from Québec's research fund FRQ (FQRSC-Team; and FQRSC-Concerted Action - Poverty and Social Exclusion-Phase 4).

## 1. Introduction

The influence of a low social class – by parental income level, or levels of education and job status – on children’s early development sets them far behind right from the start (García and Weiss, 2015). Several studies have established the early emergence of achievement gaps by race/ethnicity and socioeconomic status (SES) (Fryer and Levitt, 2004; Lee and Burkam, 2002; Princiotta and Germino-Hausken, 2006). While there is some dispute about the magnitude of these gaps and whether and how they can be explained by student characteristics and family background, there is consistent evidence that test-score gaps by race and socioeconomic status are already sizable at the end of kindergarten. Early skills gaps produce lasting consequences. If children are not ready to learn when entering Kindergarten or school, gaps are likely to persist, which in turn might determine positions at the top and bottom range of social classes (Garcia et al. 2016).

Young children from low-income families lag behind their higher-income peers (Waldfogel and Washbrook, 2011; Bradbury et al., 2012). On average, they score poorly as measured in units of a standard deviation testing on words, letters, shapes and numerals at ages four or five before entering formal school in Pre-Kindergarten or Kindergarten, and Grade 1. Bradbury, Corak, Waldfogel, and Washbrook (2015), expand the scope of previous analyses (2011, 2012). They show that significant differences in mobility-relevant skills by parental education and income are discernible as early as the age of five in all four countries examined: Australia, Canada, the United Kingdom, and the United States. These differences are generally larger in the United States and the United Kingdom than in Canada and Australia. The United States show the greatest cognitive differences by parental SES, measured by three levels of parental education.

Quality educational childcare (Li et al. 2020; Duncan and Sojourner, 2013), and pre-school and junior-kindergarten have been advocated as an effective key policy direction to close the gaps for most (Cascio, 2020; Phillips et al., 2017; Duncan and Magnuson, 2013; Cascio and Schanzenbach, 2013). Yet, socioeconomic gaps in access to quality childcare and preschool education persist (Cascio 2021; Garcia and Weiss, 2015).

Moreover, research also highlights that parenting behaviors may account for around one-third of the income-related gaps in child development (Reeves and Howard, 2014; Crawford, Goodman, and Joyce, 2011). *Early competent parenting* (nurturing, sensitivity, secure attachment with clear discipline and boundaries, emotional support, and cognitive stimulation) is positively associated with children’s early achievement and wellbeing and successful outcomes such as educational achievements (Kalil and Ryan, 2020). Closing ability gaps in the first years of life and curbing their progression as children move through school will depend also on closing *early parenting gaps*. Bradbury et al. (2015) longitudinal careful research shows that variations in parenting behavior, strongly linked to the educational SES, besides resources and income, explain about 40 percent of income-related gaps in cognitive outcomes. Parental behaviors explain more of the gap between children in the upper quantile and bottom quintile than other key factors (maternal education, family size, and race). Because the achievement gaps between low and high SES children originate prior to school entry, according to the authors, addressing the gaps

requires interventions not only in early childhood and preschool programs, but through income support programs as well as evidence-based parenting programs and preschool.

This paper contributes to the understanding of the trade-off between the spreading of universal pre-K education programs and their effectiveness in closing gaps between children coming from different socioeconomic backgrounds (Brooks-Gunn, Markman-Pithers, and Rouse, 2016), by presenting important new, non-experimental evidence on the efficacy of extensive quasi universal intervention in the early childhood years. For large representative samples of children, the evidence presented is suggestive on how preschool enrollment and duration (measured in years), before entering compulsory Grade 1, have affected their reading, math, and science literacy skills when tested in Grade 4. The motivating research question is to find out whether or not different doses of early care and education significantly influenced students' test scores. The analysis of the causal impact of 'treatment effects' makes use of a cross-national comparative approach, based on a series of international school surveys (PIRLS and TIMSS) on students' achievements conducted similarly in the same jurisdiction over years 2006 to 2019.

The paper is based on four distinctive elements. First, is the uniqueness and specific complementarities of information provided by students' parents, all related with achievement: details on the number of years of early childhood education (ECE), parental literacy and numeracy activities with children in early childhood, parents' education levels and occupation status. Second, countries that have introduced and developed more extensive universal preschool programs are selected for the empirical analysis (Nordic countries, Belgium, France) with the two most populous Canadian provinces (Ontario, Québec) representing more than half of Canada students population. All have fairly good preschool programs, and many of them with near-universal preschool for 3-5-year-olds. Third, these countries feature extended family policies: universal health insurance, income enhancing measures such as child benefits, substantial paid parental leave to stay at home after child birth, regulated working conditions, with the similar levels of quality of teaching and learning observed in schools from wealthy countries. Fourth, the statistical analysis adopts a multivalued treatment effects framework where counterfactuals are of a causal nature, taking into account early childhood education program intensities, having both care and instruction, and *gap-closing effects*.

The analysis makes three contributions. First, it measures the medium-term skills achievement of children to early preschool interventions at the end of primary school with estimated treatment effects. Second, the preschool treatment effects, expressed by three levels of preschool attendance, and in the later years surveys by programs types indicate that more is better for skills achievements, but intensities impact are generally rather small and very heterogeneous. Last, the estimations of simple but strong socioeconomic gradients (parental education and social economic status) related to students' skills achievement in the three domains (reading, math and science) reveal a more modest role than believed for preschool education as the great equalizer of social inequalities in educational achievement among primary school children for selected jurisdictions.

Section 2 discusses further the rationale for studying these questions as well as previous research results on this topic. Section 3 examines the multiyear data sets of the PIRLS and TIMSS international student achievement surveys used for estimations. Section 4 presents descriptive statistics of students' test scores, student and parental characteristics, and years of preschool enrollment in the selected countries and benchmarking Canadian provinces. Section 5 contextualizes the empirical strategy and describes the estimation methods. Section 6 summarizes the treatment effects on the treated on reading (PIRLS), math and science (TIMSS), respectively. Section 7 describes sensitivity of results to alternative duration specifications in preschool for years 2015 (PIRLS), and 2016 and 2019 (TIMSS). Section 8 details average treatment effects by preschool program types according to the ages of children. Section 9 shows differences in treatment effects by parental education. Section 10 compares three estimated social status gradients with the preschool enrollment duration gradient. Section 11 offers explanations for skills returns to preschool intensity with a discussion of policy implications. Section 12 concludes the paper with additional findings and pathways for future research analysis on this issue.

## **2. Overview of ECE studies**

Early childhood education (ECE) programs, now increasingly common in most rich countries, can be considered as requisites for four reasons. First, they provide a steady support to families with young children using childcare while they are employed or engaged in other activities. The large increase in maternal employment over the past several decades provides large evidence for their role and their strong influence on labor market behaviors (see Canada in Appendix 1). Second, the extensive research on the education component documents the importance of children's early experiences for a healthy development and later leading to better school readiness (Joo et al. 2020). Third, small scale classic early childhood experimental interventions, such as the Abecedarian Project, and the Perry Preschool Program (directed towards children living in devastating poverty and in unfavorable family background) show they may be at risk in school and later life. They provide strong evidence of the powerful long-term impacts on development of such programs (Almond and Currie, 2011; Elango et al., 2016). Because many skills are best learnt when young (Shonkoff and Phillips 2000), with their longer pay-off period, they makes such learning more productive (Garcia et al., 2016). Fourth, there may also be important "dynamic complementarities" of early learning with the acquisition of human capital at later stages (Cunha and Heckman, 2007; Heckman and Masterov, 2007; Aizer and Cunha, 2012) and the rapid brain development that preschool children experience. Providing center-based group childcare with some structured instruction or educational component for mother's young children work hours became the explicit second goal of ECE besides care. For the child development public policy orientation, there are many issues such as what, when, where, who, how (Cascio, 2021).

### **2.1 Early childcare education - American evidence**

Based on rigorous inclusion criteria over 277 studies (out of 10,309) from 1960 to 2007 of programs or interventions for children starting between birth and age 5, and the most

comprehensive meta-analytic techniques, Li et al. (2020) estimate the separate effects of the starting age, program duration, and persistence of impacts on children's cognitive<sup>1</sup> and achievement<sup>2</sup> outcomes. From studies that had to have a treatment and control/comparison group, their findings suggest effects enhancing out analyses that encompass: 1. effect sizes are larger with earlier-starting programs, those starting in infancy/toddlerhood; 2. shorter-duration programs produce larger end-of-treatment impacts than longer-duration ones; adding a year of length to a program is associated with a decrease in effect size; 3. program effects decrease as the time since the end of treatment increases; within a year of the end of treatment, but the later decline does not follow a linear manner; 4. outcomes that are more sensitive to instruction fade out more quickly than effects that are less sensitive to instruction; 5. programs targeting younger children have less fade-out impacts than those targeting older; 6. longer duration programs have more persistence effects than those with a shorter span. The authors conclude that these older interventions reflect a moderate level of effectiveness across a wide range of center-based programs, and stress the need for innovative intervention strategies to produce larger and more persistent impacts.

## 2.2 Effectiveness of preschool - American evidence

In recognition of these benefits, new generations of larger-scale preschool programs have been launched but are generally targeted. For example, the U.S. long-running Head Start programs began in the 1970s, serving low-income 3-4-year-olds, providing comprehensive early childhood education, health, nutrition, and parental involvement services, before they start elementary school. The last stage of interventions is for pre-Kindergarten (pre-K), preparing kindergarten to 4 year old children considered at-risk or disadvantaged, teaching them better behavior and providing them with early learning. Cascio survey (2021), focusing on formal programs offering group instruction for children younger than the standard eligibility age for public education, describes how ECE programs can be convincingly evaluated and why they may or may not work to narrow gaps in well-being across the lifecycle. At issue, is the optimal setting between two approaches? One is a substitutionary care for children with absent or disadvantaged parents (possibly struggling, not supportive and not stimulating), targeting low income children as for Head Start and state-funded pre-K. The other is a more classroom-based preschool program improving social and developmental skills for all children. The discussion produced multiple pre-kindergarten approaches: from part-day specialized instruction for children with special needs or immigrant children (especially those from lower socio-economic households with limited national language proficiency); to full-time and year-long programs within a public school setting under the supervision of a public school administrator, free and completely funded.

---

<sup>1</sup> This means development of knowledge, skills, problem solving and dispositions, which help children to think about and understand the world around them. Brain development is part of cognitive development.

<sup>2</sup> Letter recognition is the most obvious example, and it also includes reading, math, and numeracy. Such outcomes are more sensitive to general preschool instruction.

Studies of new generations of larger-scale targeted programs, like Head Start and pre-K, find positive effects on test scores at ages 5 to 10. But, also weakened impacts over the course of elementary school while still predicting affirmative medium-term results, such as class rank, grades, high-school completion, and also reducing incidence of behavioral problems, and health problems (Deming, 2009; Carneiro and Ginja, 2014). Studies with stringent standards of high quality (designs with pre- and post-program, randomly assigned children to a treatment or a control group) on the emerging models of universal ECE, targeted at all children, in particular pre-K, present mixed evidence on their effectiveness, with effects ranging from negative to positive, with a “fade-out” impact over school grade progression (Duncan and Magnusen, 2013; Whitehurst, 2018a, 2018b). Despite the “fadeout” effects of test score gains between groups of treated and control children, as they progress through primary school, state financed universal pre-K raises test scores, and improves other markers of school readiness that may be critical for generating long-term impacts (Cascio and Schanzenbach, 2013, 2016; Cascio, 2020, 2015).

Conditional on high-quality universal early education, with extended access for a large mix of low- and higher-income children, estimates of competence in early literacy and mathematics skills have tended to be larger for children from lower-income households or of a racial/ethnic minority. More advantaged children have also shown significant short-term positive cognitive effects (Casio, 2020; Fitzpatrick, Grissmer, and Hastedt, 2011). Program access - universal, targeted, or voluntary - with attributes such as quality teachers (those who have warmer, more positive interactions with student) and presence of higher-income children in the classroom yield different short-term positive achievement effects (Casio 2020; Lipsey, Farran, and Durkin 2018). Lipsey et al. (2018) report results where differences between treatment and control children from the third grade follow-up faded to zero and, for state test scores, turned negative as children moved through elementary school.

### **2.3 Preschool programs - European, Canadian, and Australian evidence**

Nordic and Western European countries present contrasting approaches, with similar trends in Canada, when compared to the U.S. Most countries from the European Community, including the U.K., Belgium, France, Germany and all Nordic nations, offer publicly subsidized universal childcare or provide for 3-5 year old children public school-based preschool programs aimed at promoting children’s social and cognitive development. A general concern motivating state policy intervention is to contribute to an equalization of the opportunity structures. More egalitarian income distribution is associated with generous transfers to families with children in particular. Key policies are income enhancing measures, like child benefits and high minimum wages. Larger-scale social interventions, rather than the targeted ECE programs, with nearer-universal coverage attained via heavy public subsidization of costs, have been developed over the years.

A distinct regime of much higher quality related to center-based care has been developed in several countries (notably Nordic countries) for children older than 2 years, which is ECE and pre-K offered at a low price to all families, regardless of need. Most countries across Europe were following suit, by extending public education systems to include younger children, providing



education free of charge from age 3 in general, where provision is sometimes based in the same building as primary schools (e.g. France, Belgium).<sup>3</sup>

Yet, despite enormous policy interest, evidence on the effectiveness of the extensive and more universal European childcare models and preschool programs is scarce and far from unified. There are surprisingly few nation-wide studies, all with a diversity of methods and outcomes, some focusing on cognitive development, others centered on non-cognitive skills, while some pinpoint specific aspects of child development.<sup>4</sup> In the group of studies on early schooling enrollment, most of the estimated impacts on cognitive skills are generally positive, small, unequally distributed across children from low- to high-income families, and effective mainly for disadvantaged children. An Appendix identifies empirical studies for selected countries (England, Netherlands, Spain, Italy, France, Germany, Norway, Canada and Australia) with pertinent results mostly on preschool programs.

#### **2.4 Socioeconomic gradient of children's cognitive achievement**

There is long intense interest in how differences in children's outcomes are associated with parental socioeconomic background (income, education, occupation status) as they aged from early childhood to middle childhood to adolescence. Duncan and Murnane (2011), and Ermisch, Jäntti, and Smeeding (2012) editors of different collective academic books, attribute considerable significance to SES gradient of differences in cognitive and socioemotional (non-cognitive) outcomes. Parents with better levels of education and higher incomes may invest more time, quality goods and services into their children. The resulting differences in outcomes, it is argued, emerge early at the pre-school level and then are re-enforced in childhood and the teenage years, despite the potential equalizing effect of compulsory education.

Ermisch, Jäntti, and Smeeding (2012, chap. 2) review analyses of socioeconomic gradients in children's outcomes over the life course. To harmonize parental SES measures, they employ a four-category education ranking for parents' education, using the more highly educated of the two parents or the education of single parents (some authors also use income, earnings, or occupation if available). They prefer education as a measure of SES because it is a measure of permanent income and because people with different educational qualifications face different labor markets with different rewards and opportunities and make different career path choices (e.g. artist or banker). They use these data as a separate meta-analytic database and examine how the association of child outcomes with parental SES varies across domains, countries, and child's age in an admittedly broad but informative way. Their regressions, based on 292 data points linking parental SES (as standardized by education) to various child and adult outcomes, suggest important ways in which the intergenerational gradients differ, and some in which they

---

<sup>3</sup> Regulations exist for maximum number of children allowed per staff member, training and working conditions of staff in charge of early childhood education and care, definition of appropriate curricula and governance.

<sup>4</sup> That is, the impact of preschool quality on Danish students' test results at the end of primary school's 9th grade (Bauchmüller, Gørtz, and Rasmussen, 2014); the effects on the earnings distribution of exposed children as adults of a large scale expansion of subsidized childcare in Norway (Havnes and Mogstad, 2011, 2015); a multivariate multilevel modelling of scores outcomes for Grade 4 Italian students (Grilli, Pennoni, Rampichini, and Romeo, 2016); the effects of school systems on Swedish and Norwegian students' reading performance (Rasmusson, 2016).

do not, across these dimensions. There is evidence from a number of countries that intermediate outcomes after early childhood also have a steep socioeconomic gradient.

Duncan and Murnane (2011), quote James Heckman, the Nobel economist at the University of Chicago, who argues that parenting matters as much as, if not more than, income in forming a child's cognitive ability and personality, particularly during the years before children start school.<sup>5</sup> The book's analyses identify the ways rising inequality undermines one of the most important the ability of schools to provide children with an equal chance at academic and economic success.

Parenting is multi-dimensional. Putnam (2015) describes the growing gulf between how the rich and the poor American families raise their children, while Reeves and Howard (2013) and Reeves (2017) argue that too little attention is paid to the divide affecting inherited advantage and disadvantage, the parental gap. Upper middle-class families by their income (in the upper quintile) or education (with an undergraduate degree or more) not only have better opportunities, but are also doing a lot of things right. Education is positively correlated between spouses, and, in many countries, with stability of marriage as well. More educated parents also have fewer children and have them later in life. Upper-middle class families are not only richer (with two professional incomes) and more stable, they are also more nurturing, in that they adopt more effective parenting practices and behaviors with their children. Phillips (2011) documents how parents spend different amounts and quality of time interacting with their children and exposing them to "various" environments, and how these factors can make a difference in their development and academic outcomes. They talk with their school-aged children for three hours or more per week. Phillips estimates that between birth and age six, children from high-income families will have spent 1,300 more hours in contexts other than home, school or in the care of another parent or a day-care provider than children from low-income families. In addition, the amount of time parents spend in literacy/numeracy activities is lower for low-income than for high-income families (see below for our use of such an index in the PIRLS and TIMSS surveys).

### **2.5 Universal childcare program and maternal time use trade-off**

The importance of parental time in determining child attainment has long been recognized by economists. Despite this, there are few empirical studies that analyze the effect of parental time inputs on child outcomes. Del Bono et al. (2016) using large longitudinal survey data from the UK Millennium Cohort Study, estimates the relationship between maternal time inputs and early child development. Results indicate three main results: first an overall positive relationship between two maternal time inputs (educational and recreational time) and child cognitive and emotional skills development between the ages of 3 and 7; second, these correlations are large, corresponding to 20–40 percent of the impact of having a university educated mother rather than a mother without any qualification; third, early time investments such as reading, playing,

---

<sup>5</sup> "Early life conditions and how children are stimulated play a very important role," he said. "The danger is we will revert back to the mindset of the war on poverty, when poverty was just a matter of income, and giving families more would improve the prospects of their children. If people conclude that, it's a mistake." (The New York Times, 2012/02/10/"Education Gap Grows Between Rich and Poor, Studies Say", by Sabrina Tavernise quoting Heckman).

helping with homework, providing medical care are more productive than later time investments.

Some papers have come back to older findings consistent with a psychology hypothesis (Kalil, Ryan and Corey, 2012), while others estimate negative effects of full-time maternal employment on child development (Bernal, 2008; Bernal and Keane 2011). That early high quality daycare for infants and toddlers, for parents returning to work soon after the birth of a child, may likely generate desirable effects for both cognitive and non-cognitive outcomes. Early daycare attendance positive impacts are observed for a relatively disadvantaged population, but with no effects on children from advantaged families. In particular, recent studies have examined if additional time mothers spend working associated with greater financial resources affects other time allocation decisions, especially the magnitude and types of time with children (Agostinelli and Sorrenti, 2021; Bastian and Lochner, 2020; Fort, Ichino, and Zanella, 2019). The concerns are that greater mother's labor force participation, induced by universal heavily subsidized childcare, and earned-income tax-credit, reduce parental time investments in children. Especially are single mothers, usually the main target group of these welfare programs and are considered most responsive to incentives.

These papers focus in particular on the trade-off between the income effect (economic resources) and the substitution effect (time and quality of the parent-child interactions) on a child's cognitive and behavioral development. A program that heavily subsidized daycare at age 0-4 will most likely change the amount of maternal (parental) care a child receive, depending on family context, experiencing fewer one-to-one interactions with adults. Giving the well-documented effects of subsidized childcare on maternal labor supply, attention must be given to the amount and nature of time spent, or interactions types with young children. Combining families' labor-force with time-use data sets, analyses indicate that higher wages may cause parents to substitute leisure and home production for time at work with little, or even positive, effects on time spent with children. Results depend on families' education, structure, and wages.

### **3. PIRLS and TIMSS data sets, and variables used**

International large-scale assessments (ILSAs) provide representative information about student achievement, population knowledge, skills or behaviors in certain domains, and across countries. There are some subtle conceptual differences in the skills ILSAs tests measure, with PIRLS and TIMSS focused upon "curriculum-based" measures of reading, numeracy and science literacy, while the well-known PISA surveys measure children's ability to use their skills in "real-life" situations. This analysis is based on two types of achievement for Grade 4 students (all in 4th year of primary school) aged 9-11 years, the *Progress in International Reading Literacy Study* (PIRLS) inaugurated in 2001, and the *Trends in International Mathematics and Science Study* (TIMSS), a math and science assessment conducted in fourth and eighth grade since 1995.

Our analyses are based on stratified and clustered random samples of students at the end of primary school in six selected European countries – Denmark, Finland, Norway, Sweden, France, Belgium, and in Canada represented by the two provinces of Ontario and Québec which

represent more than half of the country students' population and participated separately in all surveys. The Flemish and the French regions in Belgium also have their own school system and participated separately in the assessments. Hereafter, to simplify the presentation, no distinction will be made between country, region or province within a country; all will be designed as a jurisdiction. Norway considered that a comparison with Sweden should be considered and based on Grade 5 students (median age 10.8 years) instead of Grade 4 students (median age 9.8 years), so has two samples or only one for some years.

**Table A: Average scale score and number of Grade 4 (and Grade 5 in Norway) students with Reading, Math and Science achievement by selected jurisdiction, and year-survey from 2001 to 2019**

PIRLS Reading	2001 (Observations)		2006 (Observations)		2011 (Observations)		2016 (Observations)	
Finland	#		#		568 (4,910)		566 (4,896)	
Norway 4 <sup>th</sup>	499 (3,459)		498 (3,837)		507 (3,192)		517 (4,354)	
Norway 5 <sup>th</sup>	#		#		#		559 (4,232)	
Denmark	#		546 (4,001)		554 (4,594)		547 (3,508)	
Sweden	561 (6,044)		549 (4,394)		542 (4,707)		555 (4,525)	
Belgium Flemish	#		547 (4,479)		#		525 (5,198)	
Belgium French	#		500 (4,552)		506 (3,727)		497 (4,623)	
France	525 (3,674)		522 (4,404)		520 (4,438)		511 (4,767)	
Ontario, Canada	548 (4,295)		555 (3,988)		552 (4,561)		544 (4,270)	
Québec, Canada	537 (3,958)		533 (3,748)		538 (4,244)		547 (3,179)	
TIMSS Math, Science	2007 (Observations)		2011 (Observations)		2015 (Observations)		2019 (Observations)	
	Math	Science	Math	Science	Math	Science	Math	Science
Finland	#	#	545 (4,638)	#570	535 (5,015)	554	532 (4,730)	556
Norway 4 <sup>th</sup>	473 (4,108)	473	495 (3,121)	494	493 (4,164)	493	#	#
Norway 5 <sup>th</sup>	#	#	#	#	549 (4,329)	538	543 (3,951)	539
Denmark	523 (3,519)	523	537 (3,987)	528	539 (3,710)	527	525 (3,227)	522
Sweden	503 (4,676)	503	504 (4,663)	533	519 (8,284)	540	521 (3,965)	537
Belgium Flemish	#	#	549 (4,849)	509	546 (5,404)	517	532 (5,655)	501
Belgium French	#	#	#	#	#	#	#	#
France	#	#	#	#	488 (4,893)	487	485 (4,186)	488
Ontario, Canada	536 (3,496)	536	518 (4,570)	#528	512 (4,574)	530	512 (3,830)	524
Québec, Canada	517 (3,885)	517	533 (4,235)	516	536 (2,798)	525	543 (3,837)	522

Notes: Observations of students with test scores, where # are missing scores denoting not-participant jurisdictions. PIRLS and TIMSS scaled average, for each survey total participant countries and regions, are all 500. French Belgium schools did not participate in TIMSS surveys for Grade 4 students. In some surveys, a few other Canadian provinces occasionally participated. Number of observations for math and science are the same.

Sources: Official PIRLS and TIMSS International scores in Reading, Math and Science, various years. Results are based on each country or region five plausible values, sampling zone, replication weights, and students sampling weights.

The upper and bottom panels of Table A, identify respectively the PIRLS and the TIMSS by selected jurisdictions participating over surveys and years with number of students and their average score as calculated by the surveys statisticians. These jurisdictions have a well-established early childhood education system with high enrollment rates (see below). They therefore are well suited for the specification of both control and treatment groups for the study, because the data include retrospective information not only on whether children attended preschool, but also the number of years they spent in such settings. Section 4 below comments comparatively the international scores when converted into the z-score metric.

**3.1 PIRLS data sets** The PIRLS survey, conducted every five years since 2001, provides unique linked data on students' reading achievement, preschool attendance, home and learning environments, and parental background in several countries. Survey years are for 2001, 2006, 2011, and 2016, with a fifth cycle conducted in 2021 (results announced for, December 2022, but reported for June 2023). In 2001, 35 countries participated in the survey but few of our selected jurisdictions, while in the following surveys almost all were participants (see Table A).

Average scale scores indicate that the 2016 performance of the selected jurisdictions are all over the average scale center point (500) of all 50 countries and the eleven benchmarking participants, except French Belgium. Moreover, PIRLS provides unique data background questionnaires for students, parents, teachers, and principals for all surveys. We use the three last surveys in which almost all of our selected jurisdictions participated three times (but only twice for Finland and the Flemish Belgium region).

The PIRLS framework for reading achievement was initially developed for the first assessment in 2001, using IEA's 1991 Reading Literacy Study (Elley, 1992; Wolf 1995). Reading literacy was defined as "the ability to understand and use those written language forms required by society and/or valued by the individual." Since then, the PIRLS assessment framework has been updated for each subsequent cycle (Campbell et al., 2001; Mullis et al., 2006; Mullis et al., 2009). It focuses on two global purposes for reading that account for most of the reading done by young students, both in and out of school (with test percentages allocated to each): (a) literary experience (50%), and (b) acquire and use information (50%)(Mullis and Martin, 2015, p.13). Because young children cannot be subjected to long testing periods without suffering a loss of concentration and fatigue, the testing time is limited to 80 minutes per student, with an additional 15–30 minutes for a student questionnaire. Each student booklet consists of two 40-minute blocks of passages and items (one literary and one informational). In selecting texts for international reading literacy, potential for cultural bias has been taken into consideration.

**3.2 TIMSS data sets** The TIMSS, first administered in 1995, was the largest international student assessment study of its time evaluating students in five different grades. In the second cycle (1999), only Grade 8 students were tested. In the five following cycles (2003-2019), both 4th and 8th graders were assessed. The structure of the survey is very similar to the PIRLS; in particular the parental questionnaire introduced in 2011 is the same. In general, the Grade 4 frameworks used in TIMSS 2011, 2015 and 2019 are similar, with minor updates to particular topics in order to better reflect the curricula, standards, and frameworks of the participating countries as reported in the TIMSS 2011 Encyclopedia (Mullis et al., 2012). The *Mathematics Framework* has two assessments organized around two dimensions, where each shows the target percentage of testing time devoted to each content and cognitive domain: (a) the content dimension, specifying the subject matter to be assessed (numbers 50%, geometric shapes and measures 35%, and data display 15%); and (b) the cognitive dimension, specifying the thinking processes to be assessed (knowing 40%, applying 40%, and reasoning 20%). The *Science Framework* is organized in the same way: (a) content dimension (life science 45%, physical

science 35%, and earth science 20%), and (b) cognitive dimension with similar thinking processes. The individual student response burden for all the TIMSS' assessments is the same —that is, 72 minutes for the two assessments (36 minutes, with a break, by section), and 15-30 minutes for the student questionnaire. The bottom panel of Table A indicates students' average scaled scores in math and science by survey year for the selected jurisdictions in 2007, when only few participated, and in 2011, 2015, and 2019, used in our empirical models because of the parental questionnaire. Average scale scores show that for all year's performances the selected jurisdictions are over the average scale center point (500), except for France, and Norway younger 4th graders.

**3.3 Preschool attendance duration** The central variable, used to link preschool treatment and achievement in scores, is attendance in a preschool institution before entering Grade 1 compulsory primary education. The data sets include parental retrospective information, not only on whether children attended preschool, but also the number of years they spent in such establishment. Table B describes extensively the duration of preschool enrollment, reported by one of the parents (most of the time the mother) for the PIRLS (2006, 2011, 2016), and the TIMSS (2015, 2019) surveys. The survey asks the number of years of enrollment in preschool for a selected child, before Grade 1 (including kindergarten and not attending at all). The maximum response for years of preschool enrollment was increased to the possibility of answering that the child spent more than 3 years in 2011, and 4 years 2016, and 2019. Depending on the year and participating jurisdictions, a rather small percentage of parents (sometimes, the father or a delegate person) do not answer this question. The non-response rate is higher for Ontario and Québec, while it increased in 2016 for most countries, compared to 2006 and 2011. In our selected jurisdictions, *only a small fraction of children* never attended preschool or have been enrolled for less than one year (spending at least some months in formal care). The lengthening of average preschool attendance duration reflects a shift due to a combination of higher participation, rising intensities of childcare, and expansion of spaces offering over years, captured by the changes in survey responses choices.

For the 13 years covered by the surveys used for estimations, the group of selected European jurisdictions is observed with a large majority of students who were enrolled in ECE during their early years. The lower spells in childcare are also noticeable in the two Canadian provinces, although duration increases is clear over the years, while response rates to some questions - such as childcare enrollment - or participation rates in the parental overall questionnaire, are much lower than in the other countries. There are remarkable differences with regards to time in childcare across survey years and there is also significant cross-country variation. But the trend is clearly of a rising formal childcare duration across jurisdictions, reflecting increases in mothers' labor force participation.

**Table B: Number of students (N) and percentage distribution of preschool attendance duration in years before Grade 1 by selected jurisdiction, PIRLS 2006-2011-2016, and TIMSS 2015-2019**

Jurisdictions	Percentage of preschool duration by number of years D <sup>a</sup>						Non-		Total N
	D ≥ 3	2 < D < 3	D = 2	1 < D < 2	D ≤ 1	D = 0	Missing (%)	Missing (%)	
<b>PIRLS 2006</b>									
Denmark	78.5	14.2	2.1	1.8	2.8	0.6	3,714 (92.8)	287 (7.2)	4,001
Norway 4 <sup>th</sup>	62.6	14.6	6.5	4.3	3.5	8.5	3,525 (91.9)	312 (8.1)	3,837
Sweden	60.8	11.5	4.9	14.3	4.6	3.9	3,831 (87.2)	563 (12.8)	4,394
Belgium Fl.	85.4	11.9	1.2	0.4	0.4	0.6	4,305 (96.1)	174 (3.9)	4,479
Belgium Fr.	76.0	17.7	3.0	1.2	0.9	1.3	4,051 (89.0)	501 (11.0)	4,552
France	75.8	19.6	2.6	1.0	0.4	0.6	4,028 (91.5)	376 (8.5)	4,404
Ontario	9.4	7.7	14.3	10.6	12.6	45.4	3,457 (86.7)	531 (13.3)	3,988
Québec	11.8	7.7	12.3	10.1	24.4	33.7	3,239 (86.4)	509 (13.6)	3,748
<b>PIRLS 2011</b>									
Denmark	80.9	13.8	1.8	1.3	1.9	0.4	4,304 (93.7)	290 (6.3)	4,594
Finland	49.9	16.3	6.1	8.5	18.2	0.9	4,658 (94.9)	252 (5.1)	4,910
Norway 4 <sup>th</sup>	71.9	15.5	5.2	2.4	2.2	2.9	2,907 (91.1)	285 (8.9)	3,192
Sweden	74.3	9.9	3.0	6.6	2.4	3.9	3,992 (84.8)	715 (15.2)	4,707
Belgium Fr.	76.0	19.4	1.9	0.9	0.5	1.3	6,692 (89.8)	762 (10.2)	7,454
France	75.3	20.7	2.6	0.6	0.3	0.5	4,081 (92.0)	357 (8.0)	4,438
Ontario	18.9	12.7	35.8	12.7	15.2	4.8	3,586 (78.6)	975 (21.4)	4,561
Québec	15.3	8.3	13.4	13.2	44.5	5.3	3,653 (86.1)	591 (13.9)	4,244
<b>Percentage of preschool duration by number of years D<sup>a</sup></b>									
Jurisdictions	D ≥ 4	D = 3	D = 2	D = 1	0 < D < 1	D = 0	Non-		Total N
							Missing (%)	Missing (%)	
<b>TIMSS 2015</b>									
Denmark	54.1	39.0	4.5	1.1	0.6	0.8	3,248 (87.5)	462 (12.5)	3,710
Finland	46.8	24.0	11.6	14.8	2.0	0.8	4,724 (94.2)	291 (5.8)	5,015
Norway 4 <sup>th</sup>	74.9	19.6	3.4	0.9	0.6	0.6	1,874 (45.0)	2,289 (55.0)	4,163
Norway 5 <sup>th</sup>	71.4	21.4	4.1	1.5	0.6	1.0	1,848 (42.7)	2,481 (57.3)	4,329
Sweden	75.9	13.9	3.9	2.6	0.8	3.0	7,116 (85.9)	1,168 (14.1)	8,284
Belgium Fl.	59.8	28.1	5.0	1.8	1.3	4.0	4,827 (78.8)	577 (21.2)	5,404
France	28.3	59.0	5.5	3.1	0.8	3.4	4,144 (85.0)	729 (15.0)	4,873
Ontario	17.7	21.6	18.5	7.7	3.7	30.9	3,559 (77.8)	1,015 (22.2)	4,574
Québec	37.7	22.7	12.4	9.1	2.5	15.6	2,366 (84.6)	432 (15.4)	2,798
<b>PIRLS 2016</b>									
Denmark	83.0	12.7	2.5	0.8	0.4	0.6	3,305 (94.2)	203 (5.8)	3,508
Finland	46.7	23.6	11.9	14.9	1.8	1.1	4,540 (92.7)	356 (7.3)	4,896
Norway 5 <sup>th</sup>	68.5	14.8	3.8	7.8	2.3	2.8	4,022 (92.4)	332 (7.6)	4,354
Sweden	77.4	13.2	3.4	3.0	0.6	2.4	3,839 (84.8)	686 (15.2)	4,525
Belgium Fl.	59.8	29.3	5.1	1.7	1.1	3.0	4,615 (88.8)	583 (11.2)	5,198
Belgium Fr.		92.8 <sup>c</sup>	5.3	1.4	0.5	0.0	4,033 (87.2)	590 (12.8)	4,623
France	22.9	56.2	6.1	4.0	2.3	8.5	4,204 (88.2)	563 (11.8)	4,767
Ontario	17.6	23.5	26.5	9.5	4.1	18.8	3,354 (78.5)	916 (21.5)	4,270
Québec	36.7	10.9	14.3	9.1	2.2	17.9	2,726 (85.8)	453 (14.2)	3,179
<b>TIMSS 2019</b>									
Denmark	88.4	6.3	3.8	0.5	0.5	0.5	2,728 (59.6)	1,931 (41.4)	4,659
Finland	51.0	25.2	9.6	11.5	1.9	0.8	4,100 (86.7)	630 (13.3)	4,730
Norway 5 <sup>th</sup>	83.1	11.5	2.6	1.3	0.4	1.1	3,507 (59.9)	2,343 (40.1)	5,850
Sweden	78.2	12.2	3.2	2.1	1.0	3.3	4,369 (77.2)	1,293 (22.8)	5,662
Belgium Fl.	64.3	25.1	5.7	1.8	1.6	1.5	4,135 (88.8)	520 (11.2)	4,655
France	27.0	61.0	5.2	3.1	1.2	2.5	5,419 (88.3)	715 (11.7)	6,134
Ontario	17.0	22.8	29.4	11.5	4.3	15.0	2,506 (65.4)	1,324 (34.6)	3,830
Québec	42.3	22.6	14.5	12.8	1.9	5.9	2,848 (74.2)	989 (25.8)	3,837

Notes: a. Proportions of non-missing observations. Missing category represents mainly parental non response participation, and marginally parental non-response for preschool participation or duration. Source: Authors' computations from data sets.

### 3.4 Surveys design and test scores

International education surveys typically use a two-stage design. At stage 1, schools are randomly selected by each jurisdiction school authority (at least 150), and at stage 2, classes are randomly chosen from within each school (all students of one class only, occasionally two) in PIRLS and TIMSS. All are based on a complex psychometric design to measure skills in a number of different subject areas (reading, math, science), and within these, a number of different dimensions in order to keep the length of the test manageable for students. Participants are randomly assigned to complete particular test booklets, each including only a limited number of test questions. In both surveys, children's answers are summarized by the survey organizers into a score using an 'item-response model'. The intuition is that true abilities are unobserved, and must be estimated from the given answers. Five 'plausible values' are generated for each student, each one estimating the individual's true proficiency and using the variability among them as a measure of the imputation uncertainty, or error. The measures of tests scores have a metric that appears superficially comparable across the surveys since each year scores are scaled in both PIRLS and TIMSS by the survey organizers to have a mean across all participating countries of 500 and a standard deviation of 100. But the pool of participating jurisdictions differs across surveys and years (see above for our selected jurisdictions). As a result, jurisdiction differences in mean scores and their proper standard deviation are generated.

Although it is relatively straightforward to use proper school, class, and student weights, researchers conducting pooled analyses of multiple countries must make more difficult decisions concerning jurisdiction weights. One alternative, used here, is to leave these differences in national variances in the data sets, but to adjust the raw scores in each survey year by the mean and standard deviation in the *pooled sample* of the PIRLS (8-9 jurisdictions) and TIMSS (8 jurisdictions), where each is weighted in proportion to the size of its test sample ("house weighting").<sup>6</sup> Pooled samples scores are transformed into Z-scores national metric comparable across countries and between surveys. That is, in each survey, a student's test score is adjusted by subtracting the pooled mean and dividing by the standard deviation. With this 'international' metric, the mean score is equal to zero and the standard deviation is equal to one in the pooled sample at survey.

### 3.5 Parental socioeconomic status and family background

Each student taking the PIRLS/TIMSS assessment completes a questionnaire on aspects of home and school life, including basic demographic information, home environment, school climate for learning, and self-perception and attitudes toward reading, math and science. The language is simplified in the fourth grade version of the test and specific content is altered to be appropriate for the grade level. The response rate is very high. We did not use all the information from the surveys, except gender and age (date of birth, month, and year), and frequency of speaking tests' language at

---

<sup>6</sup> One alternative in pooling countries would be to apply 'senate weights' (weighting all jurisdictions equally) or 'student weights', where more populous countries would contribute more to the results. Only France as a country has a larger student population in all grades. Experimenting with these weights produced very similar results.



home, since other questions related to educational materials (computer, study desk, own room, internet, country-specific indicator of wealth, amount of books at home) are more present-day.

The parents (mother, father or caregiver) of each student being assessed are asked to complete a home questionnaire which requires 15-30 minutes to complete. Information on children ECE programs enrollment and number of years in this questionnaire was covered above.<sup>7</sup> It asks about home resources for literacy and numeracy, as well as early childhood literacy, numeracy, and parental activities related to science,<sup>8</sup> the child's reading and quantitative readiness when starting school, parents' attitudes toward reading and mathematics, amounts of books and books for children at home, as well as parental, both for the mother and father, education and occupation.<sup>9</sup>

**Parental socioeconomic status level** Two socioeconomic statuses were imputed for students from the parents' education levels and main job classification: education status, with three categories (lower, middle and high); and a SES index of occupational level. Table C presents the assignment rules adopted for all survey-years. The 7-8 education levels (the latest surveys distinguish master and Ph.D. levels) are reclassified into 3 levels (highest level of mother or father): 1. Upper secondary or less (the number of parents with a lower secondary education level or less is relatively small so that we do not create a particular classification for this group ). 2. Postsecondary education (more than secondary and less than university enrollment). 3. University education, undergraduate degree, master, doctorate. This variable will be our main indicator of the parental education status.<sup>10</sup> In the section for gradients estimations, the 1 to 9 categories of highest parental education (computed out of each parent's level where the highest level is attributed to the scale) are transformed in an index of years of education (based on the International Classification of Education – ISCED – as used by PISA).

The father and mother's occupation is also a widely accepted measure of socioeconomic status (SES) in sociological research, and is reliably reported in international surveys (Jerrim and Micklewright 2014). But this information is missing for 17 percent the TIMSS 2015 sample (in 2019 this percentage is higher for some jurisdictions). An index of the highest occupational status of the parents is derived from main occupational/job data for both the student's father and mother obtained from responses to open-ended questions. The responses were coded copying the PISA survey approach where four-digit ISCO codes are mapped to the international socioeconomic index of occupational status (ISEI) (Ganzeboom and Treiman, 2003). The occupational

---

<sup>7</sup> The last PIRLS 2016 survey on early childhood education also subdivided into level/type programs for children under age 3, and pre-primary education programs including kindergarten for children aged 3 or older.

<sup>8</sup> PIRLS-TIMSS Index of early literacy and numeracy activities before entering school computed elements such as: read books, tell stories, sing songs, play with abc-toys, talk about thing, talk about reading, book discussion, play word games or with shapes, board or card game, write letters or words or numbers, read aloud, read aloud signs, counting songs or things, number of toys, building blocks.

<sup>9</sup> Questions relate to if the child, mother, and father was born in country; languages spoken (or language of test) by mother and father were not systematically asked or with a similar formulation over all surveys.

<sup>10</sup> Not all parents complete the questionnaire. In the three PIRLS 2006-2016 surveys, non-response is between 16 (2006) and 12 (2016) percent. As indicated above, mean and distribution of scale scores of students for whom the parents skip their questionnaire are all significantly below the scores of students with parental information.

indexes values are used as a continuous control variable in the econometric estimations presented below. They are also transformed in terciles to reflect distribution of scores by students' SES and estimate gradients.

**Table C: Scales used to construct education levels and occupation socioeconomic statuses and indexes**

Mother or father education level	Parents' highest education level	Education status	Years index	Parents' highest occupation level	SES index
Did not go to school-1	Lower secondary	Lower	8	Has never worked for pay	22
Some primary-2	Lower secondary	Lower	10	General Laborers/Domestic worker	24
Lower secondary-3	Lower secondary	Lower	12	Skilled Fishery/agriculture worker	31
Upper secondary-4	Upper secondary	Lower	14	Plant/machine operator	33
Post-secondary-5	Post-secondary	Middle	16	Craft/trade worker/ builders	37
Short-cycle tertiary-6	Post-secondary	Middle	18	Service/Sales worker includes travel	45
Bachelor's or equivalent-7	University or higher	High	20	Clerical worker/ office clerks	49
Master's or equivalent-8	University or higher	High	21	Technician/Associate Professional	52
Doctor or equivalent-9	University or higher	High	23	Small Business owner	57
				Corporate Manager/Senior official	67
				Professional includes scientists	73
Not applicable	Not applicable	NA	NA	Not applicable	NA

Source: PIRLS and TIMSS surveys parental questionnaire and authors' constructed indexes.

**Family background covariates variables** Statistics of all variables used by survey-year are presented in Table D.1.1 (PIRLS) and D.1.2 (TIMSS). Family background information, in particular parental education and occupation, reported by one of the parents in the parental survey questionnaire, relates to the time of the survey. When the student was younger during preschool years, these characteristics could have been different, but admit they are very good proxies. The only other variables, which bear on the preschool years of the student, are preschool attendance and duration, and indexes of early literacy and numeracy activities before entering school (footnote 8 for a description) constructed by the survey. Since the student preschool years are not very distant in time, and that preschool attendance-duration is a significant life fact for the family, the analysis assumes parental information matches up with reality.

The treatment effects estimations framework (see section 5 for the econometric methodology) combines two models (*treatment status* and *outcomes equations*) that must be specified. Five variables available appearing in all surveys are retained as covariates in the *treatment-assignment model*: 1. Sex and number of month-age of the student; 2. Parental education levels, highest level reached by either father or mother, whichever is highest (low, middle, high); 3. Occupational socioeconomic status of the family, declared for either father or mother, whichever is at highest level (treated as a continuous index variable); 4. A dummy for language of the test spoken at home: frequently versus sometimes or never, or when available, the child's immigration status; 5. Year dummy variable for estimation with pooled data sets.

For the *outcome model* equation, eight variables in all the surveys are used: 1. Outcome of interest are the z-scores (reading, math or science as the dependents variables). 2. Sex of the student. 3. Age-month of the student (treated as a continuous variable). 4. Parental education

levels, highest level reached by either father or mother, whichever is highest (low, medium, high). 5. A dummy for language of the test spoken at home: frequently versus sometimes or never, or when available, the child's immigration status. A dummy for the language usually used at home, same as test versus sometimes or never. 6. Occupational socio-economic status of the family (treated as a continuous variable). 7. Early preschool literacy and numeracy parental activities before school (treated as a continuous variable, except in 2006 available as categories).<sup>11</sup> 8. Year dummy variable for estimation with pooled survey data sets.

The preschool treatment variable represents both preschool enrollment and duration. It contains three integer values representing the duration levels: 1. A low level of attendance is 1 year or less (including no attendance as declared by the parent); 2. A medium level attendance is more than 1 year and less than 3 years; 3. A high level of attendance is 3 years or more.

---

<sup>11</sup> The surveys statisticians have also constructed an index of home resources for learning, composed of parental education and occupation, number of books in the home (from parents, from students), access to internet and own room (from students). This index largely correlated to parental characteristics was not used.

**Table D1: Descriptive statistics of variables used for estimation by jurisdiction, PIRLS 2006, 2011, 2016**

Jurisdiction	Denmark	Finland	Norway 4 <sup>th</sup>	Norway 5 <sup>th</sup>	Sweden	France	Belgium Flemish	Belgium French	Ontario	Québec
Obs. 2006	3,714	#	3,525	#	3,831	4,028	4,305	4,051	3,457	3,239
Obs. 2011	4,304	4,658	2,907	#	3,992	4,081	#	6,692	3,586	3,653
Obs. 2016	3,305	4,540	4,022	3,933	3,839	4,204	4,615	4,033	3,354	2,726
Male student (percentage)										
2006	48.2	#	50.0	#	51.2	51.2	50.6	48.9	49.6	49.4
2011	49.2	50.5	48.6	#	50.4	50.0	#	50.7	49.4	50.0
2016	48.4	49.7	49.6	49.1	48.6	48.8	48.5	50.4	48.5	47.4
Student's age (years)										
Mean 2006	10.9	#	9.8	#	10.8	10.0	10.0	9.9	9.8	10.1
Mean 2011	10.8	10.8	9.7	#	10.7	9.9	#	10.0	9.8	10.1
Mean 2016	10.8	10.8	9.8	10.8	10.7	9.7	10.0	9.9	9.8	10.1
Early parental literacy activities before school (discrete variable: high level percentage)										
2006	53.0	#	49.7	#	47.2	59.0	40.9	52.9	70.5	64.9
Early parental literacy activities before school (continuous variable)										
Mean 2011	9.9	9.8	10.1	#	9.9	10.0	#	9.7	10.8	10.3
Mean 2016	10.0	9.9	10.1	10.2	10.0	10.0	9.5	9.7	10.8	10.3
Immigration status: percentage born outside country										
2006	4.7	#	4.1	#	5.4	5.1	4.7	9.7	9.0	6.9
2016	4.6	4.1	6.4	6.8	6.8	5.2	5.7	9.2	9.8	9.3
Language of test used at home: percentage sometimes or never										
2006	17.8	#	17.8	#	26.5	33.7	21.9	36.2	52.5	41.5
2011	18.1	11.1	19.4	#	21.9	22.8	#	29.0	44.9	30.6
2016	10.6	9.4	13.8	11.6	13.4	18.2	23.2	22.6	38.5	30.0
Index of occupational status - SES (mean continuous variable)										
Mean 2006	55.1	#	59.2	#	57.8	53.6	54.7	54.0	58.0	57.5
Mean 2011	55.9	59.7	62.0	#	58.2	54.0	#	54.6	59.5	58.9
Mean 2016	57.6	61.5	63.1	63.3	60.5	55.2	57.0	56.3	61.8	61.5
Parental education levels (percentage)										
Low 2006	20.0	#	22.1	#	27.9	56.8	41.1	41.0	18.1	21.3
Med. 2006	28.5	#	29.3	#	35.7	15.6	26.9	48.8	36.5	37.5
High 2006	51.5	#	48.6	#	36.4	27.6	32.0	10.2	45.3	41.2
Low 2011	17.8	27.6	16.9	#	29.9	50.4	#	35.1	13.9	14.1
Med. 2011	24.6	27.1	22.3	#	27.3	17.9	#	12.7	33.7	38.8
High 2011	57.6	45.3	60.8	#	42.8	31.7	#	52.2	52.4	47.1
Low 2016	12.8	24.4	13.5	13.9	20.2	44.0	21.3	28.4	9.5	9.0
Med. 2016	23.2	20.1	21.8	22.0	25.7	20.2	19.9	12.5	34.2	36.5
High 2016	64.0	55.5	64.7	64.1	54.1	35.8	58.8	59.1	56.3	54.5
Preschool attendance levels (percentage)										
Low 2006	3.4	#	12.0	#	8.5	1.0	1.0	2.2	58.0	58.1
Med. 2006	18.1	#	25.4	#	30.7	23.2	13.5	21.8	32.6	30.1
High 2006	78.5	#	62.6	#	60.8	75.8	85.5	76.0	9.4	11.8
Low 2011	2.2	19.1	5.0	#	6.2	0.8	#	1.8	20.0	49.7
Med. 2011	16.9	30.9	23.1	#	19.5	23.8	#	22.2	61.1	35.0
High 2011	80.9	50.0	71.9	#	74.3	75.4	#	76.0	18.9	15.3
Low 2016	1.8	17.8	12.9	13.2	5.7	14.4	5.3	1.9	27.2	18.5
Med. 2016	2.5	11.9	3.8	4.5	3.5	6.1	5.2	5.3	28.1	15.5
High 2016	95.7	70.3	83.3	82.3	90.8	79.3	89.5	92.8	44.7	66.0
Participation in early age childhood and educational program(s) (percentage 2016)										
No program	0.4	1.3	1.6	2.1	2.5	8.2	2.2	#	20.1	19.3
Early age	0.0	1.6	1.9	1.7	2.0	3.2	0.5	#	4.3	4.7
Pre-primary	35.5	47.6	5.2	6.3	7.7	49.8	29.9	#	36.6	38.8
Both program	64.1	49.5	91.3	89.9	87.8	38.8	67.4	#	39.0	37.2

Notes and sources: See End of Table D2.

**Table D2: Descriptive statistics of variables used for estimation by jurisdiction, TIMSS 2015 and 2019**

Jurisdiction	Denmark	Finland	Norway 4 <sup>th</sup>	Norway 5 <sup>th</sup>	Sweden	France	Belgium Flemish	Ontario	Québec
Obs. 2015	3,248	4,724	1,865	1,848	7,116	4,144	4,827	3,559	2,366
Obs. 2019	2,776	5,938	#	3,507	4,618	5,375	4,220	2,317	2,220
Male student (percentage)									
2015	49.8	50.7	51.7	49.3	48.4	49.1	48.8	50.0	48.5
2019	49.5	50.1	#	50.7	48.8	50.3	48.9	49.4	50.2
Student's age (years)									
Mean 2015	10.9	10.8	9.7	10.7	10.8	9.9	10.1	9.8	10.1
Mean 2019	10.8	10.8	#	10.7	10.8	9.9	10.0	9.8	10.1
Early parental numeracy activities before school (continuous variable)									
Mean 2015	9.8	9.6	10.1	10.1	9.5	10.2	9.7	10.7	10.4
Mean 2019	10.2	9.9	#	10.2	9.7	10.6	9.8	11.2	10.8
Immigration status: born outside country (percentage)									
2015	4.1	4.3	5.7	7.2	7.0	5.1	6.0	11.5	9.0
2019	5.1	4.0	#	6.8	9.9	5.6	7.0	12.4	10.2
Language of test used at home: sometimes or never (percentage)									
2015	11.0	10.3	13.5	12.2	14.8	19.0	20.7	25.3	25.4
2019	11.0	10.6	#	10.2	17.5	19.1	23.6	33.3	27.0
Index of occupational status - SES (continuous variable)									
Mean 2015	57.6	60.1	63.6	62.9	60.7	53.7	55.2	59.2	60.5
Mean 2019	62.8	62.2	##	##	60.6	56.1	56.2	61.7	62.0
Parental education levels (percentage)									
Low 2015	10.7	26.9	12.5	12.5	21.3	45.5	23.6	13.0	10.3
Med. 2015	26.1	21.2	24.3	25.6	24.7	21.1	19.8	32.7	32.7
High 2015	63.2	51.9	63.2	61.9	54.0	33.4	56.6	54.3	57.0
Low 2019	9.7	23.6	#	11.4	20.1	38.1	21.2	8.8	8.9
Med. 2019	17.2	14.3	#	19.0	23.0	21.0	18.2	30.2	33.2
High 2019	73.1	62.1	#	69.6	56.9	40.9	60.6	61.0	57.9
Preschool attendance levels (percentage)									
Low 2015	2.5	17.7	2.1	3.1	6.4	7.3	7.0	42.3	27.2
Med. 2015	4.5	11.6	3.1	4.1	3.9	5.5	5.0	18.5	12.4
High 2015	93.0	70.8	94.8	92.8	89.8	87.3	87.9	39.2	60.4
Low 2019	1.5	14.2	#	2.8	6.3	6.8	4.9	30.8	20.6
Med. 2019	3.8	9.6	#	2.6	3.2	5.2	5.7	29.4	14.5
High 2019	94.7	76.2	#	94.6	90.5	88.0	89.4	39.8	64.9
Participation in early age childhood and educational program(s) (percentage 2015)									
No program	0.7	0.8	0.7	1.0	2.9	2.5	2.2	32.5	15.1
Early age	1.9	8.0	21.3	19.2	18.5	7.9	3.6	13.8	7.5
Pre-primary	47.9	51.4	14.4	17.7	10.0	55.5	35.0	32.8	47.9
Both program	49.6	39.9	63.6	62.1	68.6	34.1	59.2	20.9	29.6
Participation in early age childhood and educational program(s) (percentage 2019)									
No program	0.5	1.0	#	1.5	3.9	2.8	1.7	15.9	6.2
Early age	3.0	1.4	#	9.3	2.9	2.5	0.8	7.3	11.7
Pre-primary	5.2	43.6	#	8.0	6.5	54.9	25.0	38.7	46.8
Both program	91.3	54.0	#	81.2	86.8	39.7	72.5	38.0	35.3

Notes: Statistics calculated for the observations with no missing value on the preschool attendance duration variable. Missing values of the other variables are ignored. Parental education levels: Lower (upper secondary and less), medium (post-secondary but not university) and high (Bachelor's level or higher). Preschool attendance levels: Lower (up to one year, including no attendance), medium (between one and three years), and high (three years and more). # indicates non-participation of the jurisdiction or group of students to the survey; ## indicates a question not administered in the survey. Sources: Author's computations from PIRLS (2006, 2011, 2016) and TIMSS (2015, 2019) data sets.

## 4. Descriptive statistical analysis

### 4.1 Reading (PIRLS) 2006-2011-2016

Left-panel in Table 1 shows cross-country differences, in the mean and percentile distribution of Grade 4 students' overall international reading scores (converted into the z-score metric) by selected jurisdiction (countries and benchmarking provinces), for the three PIRLS surveys used for estimation (2006, 2011, and 2016). A first observation is that, in all three survey years, there are large cross-country disparities in the mean and median z-scores compared to the calculated overall "national" values (shown in Table A). Moreover, over the years 2006 or 2011 to 2016, in some jurisdictions there are few changes in these mean z-scores (Denmark, Sweden, Finland, Belgium French, while in other some changes are equal to 0.16 to 0.25 of a standard-deviation (SD) with progress in Norway Grade 4 and Québec, and a fall in France and Ontario). In some cases, non-participating jurisdictions as well as Norway partial survey by grade could have tainted the annual statistics. This highlights two important points. First, France, and French Belgium (Flemish Belgium participated in two of the three surveys) children's reading proficiency is lagging behind that of Denmark, Finland, Norway Grade 5 present only in 2016, Sweden, Ontario, and Québec. There is no catching-up from 2006 to 2016 except for Norway Grade 4 (-0.498 to -0.236). For instance, mean reading achievement in Norway Grade 4, France and French Belgium is 0.40-0.75 SD lower than Denmark and Sweden in 2006, 0.60-0.75 SD lower in 2016, and more than 0.50 SD behind the top performer (Finland). Second, there is clear evidence that skills performance at different percentile points of the score distribution grow systemically for each region and year (from P10 to P90). In addition the 75th and 25th percentile gaps (P75-P25), a relatively plain disparity measure, presented in the bottom line of each panel in Table 1, indicate no shrinking gap in children's educational outcomes for each region over the period 2006-2016. On the one hand, this suggests that primary schools are facing similar skills gap in reading, measured in Grade 4, for each new cohorts in schools from 2006 to 2016. On the other hand, skill gaps could have been higher in Grade 1.

Table 1 also highlights some other noteworthy differences. Notably, students in jurisdictions performing well above the international average near the end of primary school generally manage to maintain their strong performance over the years. The same is also true at the other end of the spectrum, where jurisdictions with a large proportion of students (indicated by the N% statistic) performing poorly in Grade 4 relatively to the international average tend to perform poorly over the years, even though preschool attendance increases (see following statistics). France is an example of low performing countries. In other words, once a country falls behind in the educational achievement race, it seems difficult to catch up.

To gain further insight into this issue of each skill performance for the highest and lowest achievers, observed preschool attendances as well as students socioeconomic statuses differences in achievement must be considered. In the Statistical Annex, Table A1 presents for the selected jurisdictions for PIRLS-TIMSS (2006, 2011, 2016; 2015, 2019, respectively), descriptive statistics of the number of 4th graders, by the percentile distribution of three levels, of preschool duration attendance in years classified for the three levels of parental education.

The same levels are used in the econometric estimations. It is easy to note that, for all jurisdictions, higher preschool duration (3 year or more) is associated with a higher parental education level, while for each year survey this proportion stepped-up. Table A1 indicates that the proportions of children in each preschool intensity for a given parental education level have reversed: over time, lower and middle preschool attendances are more the choices of higher education parents; their children are concentrated in higher attendance level; children of lower educated parents represent a leading share of the lower and medium attendance levels.

Table A2 shows the percentiles of reading z-scores desegregated by the three levels of preschool duration (left panel), and by the three parental education levels (right panel). The 75th and 25th percentile gaps (P75-P25) are also computed for each jurisdiction. As documented in Table B, cross-jurisdiction differences in preschool intensity are highly noticeable, even if over time, the proportion of students attending middle or high levels dragged in all regions, with the two groups away from the lower attendance level. Parents in Finland, Ontario, Québec, have on average enrolled less children in childcare.

The left panel of Table A2 (percentile attendance levels) may give a mitigated picture of the preschool attendance changes, and suggest that mean score performance is similar for a low or high preschool attendance, while both increase over the distribution implying less students' performance for a middle pattern. But trends are much sharper, with a high preschool intensity associated with a higher mean, or a less negative one for low-achievement regions (Norway grade 4, French Belgium, and France). Another interesting observation is the importance of the lower and medium preschool intensities for some regions (Finland, Ontario, Québec, and to a lesser degree Sweden) where mean z-scores remain positive and show significant achievement, along the distribution of scores.

The same comparison by jurisdiction from the right panels of the Table A2, for scores and parental education levels, strongly suggests that the inequality of children's educational opportunities is strongly associated with the family background education level. The same patterns prevail over the percentile distribution of scores. On the basis of the mean and percentile points, the differences in test scores between children whose parental education is in the higher category versus those whose family is in the lower category, are accentuated for all three surveys. While the ratios in the lower preschool category slip to become less significant, the proportion of children of parent(s) with a higher education level has substantially increased over the years.

Finally, Table A5 presents the number of 4th graders -and 5th in Norway-, the mean and the percentile distribution of their reading z-scores by the occupational index of their parents (see Table C) transformed in terciles. The cross-jurisdiction differences between the lowest and the highest terciles show the existence of a sizeable socio-economic gradient in students' reading skills. Compared to the other two measures of the students' background, terciles levels display very similar patterns (means, percentile points, and the P75-P25 gaps). The robustness of these associations is considered below for more formal gradients estimates.

## 4.2 Math and Science (TIMSS) 2015-2019

The same descriptive statistics are computed with the 2015 and 2019 surveys data sets. Survey 2011 is skipped because information on preschool attendance and parental characteristics are available for only four jurisdictions (Finland, Norway, Sweden, and Québec<sup>12</sup>). Cross-country differences, and scores converted into the international z-score metric are presented in the right-side panel of Table 1 and in Tables A3-A4 (left panel for math and right panel for science). Table 1, for the ten jurisdictions, indicates for math and science z-scores, rather large cross-country differences as for the PIRLS reading z-scores. They follow rather similar patterns to those in the PIRLS surveys, but with some changes. Denmark, Finland, and Flemish Belgium students are well above the international average in math, with France and Québec a little over the average and all other regions are observed with a negative z-score mean. Students in Finland, Sweden and Ontario stand out with high averages. The statistics make it clear that there is a score progression with the percentile points on the distribution of skill achievements, while the 75th and 25th percentile gaps document a 0.10 to 0.30 standard deviation difference in scores.

Tables A1, A3-A4, and A6 in the Annex present the association between students' educational outcomes and the three family background measures by jurisdiction, used above for TIMSS 2011 and 2015 respectively. The shifts between attendance and education levels towards more emphasis by parents of higher attendances, as well as the decreases in lower attendances by less educated parents are clear in Table A1 for all jurisdictions. There are strong connections between preschool intensity levels and mean students' z-scores, for both tests and both surveys (Table A3). They are clearly underlined, except for Finland and Québec in 2015. Statistics for 2015 and 2019 also show the presence of large gradients along the percentile distribution of scores as well as an extended bandwidth for the 75th and 25th percentile gaps. The percentile distribution of students' overall math and science z-scores by region, and parental education in Table A4 based on both surveys, reiterate the same basic trends of a continued rise of performance measures when parental education is higher. Finally, Tables A6 showing cross-jurisdiction differences in z-scores between the lowest and the highest terciles of socioeconomic status, indicate sizeable gradients in students' math and science skills attached to parental occupation levels, computed in weighted terciles. For math and science, z-scores attainments unfold as if parental backgrounds take over more influence that preschool attendance intensity, abstracting from the impacts of schools through ignorance of current quality (even though the information described by students' teachers on their qualifications, experience, and teaching methods).

## 5. Empirical modelling strategy

**5.1 Causality and multivalued treatment** Critical challenge, to evaluate treatment effects, is to avoid the recurrent problem of endogeneity. Because of the fundamental problem of causal inference, unit-level causal effects cannot be directly observed. Strategies adopting more complex econometric strategies, with a diversity of plausible counterfactuals, can help

---

<sup>12</sup> In these jurisdictions, the same students were participants in the PIRLS and TIMSS 2011 surveys.



identify causal relationships. To catch up a dearth of studies on the longer-term effects of pre-K, Cascio (2021) suggests using credible identifying variation in the proposed mechanism itself of interventions to later-life well-being. Cordero, Cristóbal, and Santín (2018) provide a similar overview of the challenge using quasi experimental techniques applied to three international large-scale comparative education assessments, PISA, PIRLS, and TIMSS, over the period 2004-2016. However, exogenous variations in databases have to be identified in order to apply causal inference techniques.

Most papers examining childcare reforms focus on intention-to-treat effects, partly because information on individual childcare attendance duration is unavailable (see, e.g., Baker, Gruber, and Milligan 2008; Lefebvre and Merrigan 2008; Havnes and Mogstad 2011, 2015; Felfe, Nollenberger, and Rodriguez-Planas 2015).<sup>13</sup> Without information on individual treatment status, however, it is impossible to determine whether heterogeneity in intention-to-treat effects is caused by the differential take-up of children or by heterogeneous responses to childcare attendance. For example, larger intention-to-treat effects at the bottom or middle part of the outcome distribution found by Havnes and Mogstad (2015) may either be produced by differences in childcare take up, or by differences in the impacts of uptake.

**5.2 Multivalued treatment models** Nonetheless, considering the uniqueness of the data sets on preschool duration for many countries, convincing type of modelling can be adopted. The fundamental pillar of the counterfactual theory of causation with respect to the estimation of treatment effects based on observational data was developed by Rubin (1974). Seminal work by Imbens (2000, 2004) and Lechner (2001) have subsequently given rise to several methods by extending Rosenbaum and Rubin's (1983) propensity score framework for binary treatments to multivalued treatments. Uysal (2015), with Linden et al. (2016) generalize the Hirano and Imbens (2001) approach to binary treatment to multivalued treatment effects estimators.

Two types of estimation methods for treatment effects of preschool intensities are used. The first one is an inverse probability weighted regression adjustment estimator. This weighted least square doubly robust (WLS-DR) estimator, is a 3-steps estimation modeling: Step 1, uses a multinomial logit of preschool intensity choices by parents, which is function of the family context variables, for the treatment model; Step 2, calls a propensity score weighting by the inverse of the estimated probabilities that children receive a specific treatment out of three to correct for missing data on the potential outcomes; Step 3, ends with a weighted least squares doubly-robust estimates, which are function of the covariate variables and the preceding parameters, for the outcome(s) model.

The second method uses a one-step generalized method of moments (GMM) approach for the estimation of the outcomes equations using treatment assignment estimates and inverse probability of treatment weighting (IPTW) regression adjusted (RA). As for WLS-DR, only one of the two estimations (treatment status, and outcomes) must be correctly specified for the IPTW-

---

<sup>13</sup> In Haec, Lefebvre, and Merrigan (2015) the estimates of the impact of Québec's low-fee policy are based on hours in formal subsidized childcare for a series of young children cohorts and number of weeks worked by the mothers who mostly use subsidized formal childcare.

RA estimator to be consistent. Estimates based on this approach conducted using Stata version 16.0 are too similar with those from WLS-DR to be presented separately.<sup>14</sup> The two main assumptions of both approaches, supposed respected, are weak un-confoundedness in the choice of covariates selected for inclusion in both the treatment and outcomes equations, and sufficient overlap.

The Jackknife Repeated Replication (JRR) procedure to estimate sampling variances was chosen because it is computationally straightforward and provides approximately unbiased estimates of the sampling variances of regression coefficients and sampling errors of means, totals, and percentages. JRR incorporates in its repeated draws of subsamples, the stratification of schools and the clustering of students within schools. Students' scores are presented with five plausible values. To use 'correctly' these values, the survey organizers suggest that one should follow a version of 'Rubin's rules' for handling multiple imputations (OECD 2009; Rubin 1987).<sup>15</sup>

**5.3 Counterfactuals and Average treatment of the treated (ATETs)** When a treatment is binary, each person in a selected sample could either "receive the treatment" or "did not participate or receive the treatment", which opens up for observational data a variety of counterfactual frameworks such as the quasi experimental techniques mentioned above as a replication. The causal effect of such a course could very well differ from person to person (this is referred to as treatment heterogeneity). When quantitative measures vary across observational units a canonical summary statistic is the average. The Average Treatment Effect (ATE) is simply the average of the individual treatment effects of the population under consideration. A simple counterfactual is the Average Treatment Effect of the Treated (ATET), the average of the individual treatment effects of those treated (hence a subsample of the control group).

Multivalued treatments refer to a world in which each person may receive one of several different treatments or else not receive any of them. The concern in this case is to know how effective each of those programs relative adjacent treatments (including not participating). In our setting, before entering Grade 1, a child could have attended three different intensity of preschool educational care: (a) no attendance at all or for a few months to 1 year, the low level (or L); (b) for more than 1 year but less than 3, the middle-level (or M); and (c) for 3 years or more, the high level (or H). The high duration level of attendance ("services") implies that a child has been cared for by someone else's home or at home, most of the time.

The effects of randomly assigning a person with observed family and child characteristics  $X$  to each undertaking, defined as the average treatment effect (ATE) can be easily computed for each

---

<sup>14</sup> Linden et al. (2016), in a Monte Carlo study, examine the performance of a 'doubly-robust' estimator (WLS-DR) that models the treatment assignment and the outcome variable within the same framework using an inverse probability of treatment weighting (IPTW) combined with a conventional regression adjustment (IPTW-RA). The relative efficiency of commonly used regression-based methods, where multivalued treatment effects are estimated with the un-confoundedness assumption in pretest–posttest studies, was similar.

<sup>15</sup> This procedure can be divided into two steps: Step 1: Estimate the statistic/model of interest five times, using each of the plausible values. This will generate five separate parameter estimates ( $\beta_{pv}$ ) and five estimates of the sampling error ( $\sigma_{pv}$ ). Step 2: To produce the final parameter and sampling error estimates, take the average of the five estimates produced in step 1. Note that the JRR weights are also applied whenever the model is estimated.

of the three treatment groups. Though, we would rather know how much the high (or middle) treatment increases scores among the children who actually have received a low-service (or middle) level. To answer that question, we must examine the average treatment effect among the treated (ATET) rather than the ATE.

Defining the concept of ATET is more complicated. Multivalued treatments increase the number of parameters that must be estimated and makes the notation more elaborated and their interpretation less direct. First, the control groups (L, M, and H) to be considered as the basis for a binary comparison (Middle versus Low; High versus Low; High versus Middle) must be specified. Second to compute the ATETs, we must specify two treatment levels: (i) the actual or observed treatment level received by the subjects we are interested; and, (ii) the predicted treatment level for subjects who do not attend the treatment level we want to compare them with.

More formally:

1. For the entire group of children, the ATEs are  $T_{jk} = E(y_{ij} - y_{ik})$ ,  $j, k = \{L, M, H\}$ , where  $y_{ij}$  and  $y_{ik}$  are the observed or predicted scores of child  $i$  given
2. For the sub-group of children whose parent chose the attendance level  $j$ , the preschool attendance levels are  $j$  and  $k$  respectively. The ATETs are  $T_{jk|j} = E[(y_{ij} - y_{ik})|A_i = j]$ ,  $j, k = \{L, M, H\}$ , where the observed score are  $y_{ij}$  the predicted scores of child  $i$  given the preschool attendance levels  $j$  and  $k$  respectively and  $y_{ik}$  the attendance level observed.
3. Finally, for the sub-group of children whose parents chose the attendance level  $k$ , the ATETs are  $T_{jk|k} = E[(y_{ij} - y_{ik})|A_i = k]$ ,  $j, k = \{L, M, H\}$ , where the predicted scores  $y_{ij}$  are, and  $y_{ik}$  the observed scores of child  $i$  given the preschool attendance levels  $j$  and  $k$  respectively and  $A_i$  is the attendance level observed.

Among the many ATETs possible, our attention will be focused on: (1) the effects on children's scores of parents choosing the middle instead of the lower level of preschool attendance for children whose parents chose the middle ( $T_{M|L}$ ) or the lower ( $T_{M|L}$ ) level; (2) the effects of choosing the higher instead of the lower level of preschool attendance for children whose parents chose the higher ( $T_{H|H}$ ) or the lower ( $T_{H|L}$ ) level; and finally (3) the effects of choosing the higher instead of the middle level of preschool attendance for children whose parents chose the higher ( $T_{H|M}$ ) or the middle ( $T_{H|M}$ ) level

There are two components for the ATETs: the selection bias (i.e., the mean bias of selection on observed characteristics in the absence of attending a higher duration) and the sorting gain (i.e., the average additional duration gain for students who attend a higher duration relative to that for students who do not attend such a duration. Some changes in estimated effects presented in the next section are more pronounced for the comparison of higher levels with small (no or small duration treatment). For example, if  $T_{21|2}$  is greater than  $T_{21|1}$ , the treatment effect is "efficient" in terms of the enrollment of children to the particular treatment level 2, that is, the children who would gain more for the treatment level 2 if enrolled in this childcare service. The difference

between  $T_{21|2}$  and  $T_{21|1}$  is the sorting 'gain' (Heckman and Li 2004). Considering treatment 2 as a return over treatment 1 (low childcare intensity), positive sorting gains would imply that the children who would benefit more from a higher childcare intensity, are cared appropriately when cared more in preschool institution. Meanwhile, negative sorting gains would indicate that there may be children with a lower level of childcare, who could have benefited from a higher duration of childcare. All possible pairwise comparisons for the three levels of preschool childcare are considered.

## 6. Average treatment effects of the treated (ATET)

Table 2 presents average treatment effects of the treated, regrouping jurisdiction by geography, tests z-scores (reading, math, science), and survey-year. The changes in the treatment levels are M (middle) versus L (low) ( $T_{ML}$ ), H (high) versus L (low) ( $T_{HL}$ ), and H (high) versus M (middle) ( $T_{HM}$ ), estimated for each control group of children whose duration of preschool attendance corresponds to one of the treatments compared. The reported estimates tests changes are measured in units of standard deviation (referred to as SD units). Besides the standard errors, the levels of significance of the estimates are reported based on p-values whose calculation takes the clustering of children within schools into account. Estimations are conducted with annual survey data for each jurisdiction.

To increase the accuracy of estimations, pooled samples over the years were also used whenever allowed by pooling tests. In the specification of the model estimated with WLS-DR, coefficients measuring treatment effects interacted with time dummies to capture their changes over time.

Before discussing the results for specific test scores, and in the following section for other duration specifications, two general observations can be formulated. First, there are a large number of estimated treatment effects with a plurality not significant. This is not surprising considering the diversity of achievements and expanse years of survey. When significant, the effects are in general for reading and science. As expected, most effects are positive but some are negative and significant, especially for the older surveys and the lower duration. They are larger when associated with longer childcare durations. A second consideration is that proportions of children enrolled in childcare before entering Grade 1 and, consequently, average duration of childcare, have increased over the years for all jurisdictions. Thus, estimating the impact of childcare duration on test scores without taking the endogeneity of this variable into account, as the models do, would possibly induce an omitted "family-ability bias" since families where parental education and/or occupational status is higher take more advantage of childcare. Also, this decreases the gaps between the intensities of preschool across children over the top and the bottom range of social class, as shown in the descriptive statistics.

### 6.1 Reading ATETs 2006, 2011, and 2016

#### 6.1.1 Denmark, Finland, Norway, Sweden

The reading ATETs estimates of first jurisdiction presented in Table 2 are for Denmark students who have very good scores in reading over the years (Table 1). These children were

enrolled intensively and increasingly in childcare. Treatment effects estimated with pooled data are strongly positive and significant (0.22-0.11 SD for  $T_{HM}$  and 0.39-0.32 SD for  $T_{HL}$ ). Moreover, since  $T_{ML}$  and  $T_{HM}$  must add up roughly to  $T_{HL}$ , the estimated treatment effects of  $T_{ML}$  (0.19-0.13 SD), although not significant, are highly credible. However, treatment effects estimated with data for 2006 are mostly negative and significant, except for the higher duration versus the middle.

Finnish students also have high achievement results in reading, and when younger were enrolled very moderately in terms of childcare intensities. Estimated treatments  $T_{ML}$  and  $T_{HL}$  are significantly negative (-0.10 to -0.07 SD) suggesting that long enrollment in childcare does not generate gains in this country. Pooled effect for higher duration versus the middle does not indicate is not significant.

Norway students deserve special attention insofar as students in Grade 4 are on average one year younger than students in all other regions selected for the study, and have lower z-scores. Norway suggests that Grade 5 students should be compared with the other countries' Grade 4 students. But on average, students in Grade 5 have higher scores than all the other countries' Grade 4 students. As children, they attend childcare rather intensively, but the estimates do not capture its effects over the years, except for treatment  $T_{HM}$  (0.17-0.08 SD) when measured with pooled data.

Swedish students are much similar in terms of reading achievement and childcare patterns over the years. Significant treatment effects are measured, over the years and with pooled data, for  $T_{HM}$  (0.08-0.07 SD) and  $T_{HL}$  (0.16-0.13 SD). Attending childcare between one and three years instead of one year or less has no effect on z-scores except for children with the lowest level of attendance when measured with pooled data (2006-2016).

### **6.1.2 Flemish Belgium, French Belgium, France**

These three jurisdictions appear to have developed similar programs of educational childcare. While the two Belgian regions have observable differences in achievement in reading (z-scores), their enrollment childcare intensities are quite similar with 95 % and 98 % of children attending childcare for 3 years or more in Flemish and French Belgium respectively.

For Flemish students, very significant treatment effects are for  $T_{HM}$  estimated between (0.09-0.14 SD), and (0.14-0.30 SD) for  $T_{HL}$ . French Belgium has very low achievement scores compared to other jurisdictions and the rather intensive childcare duration over the years does result in robust treatment effects for  $T_{HM}$  (0.11-0.11 SD), and a large effect of (0.26 SD) for  $T_{HL|L}$  in year 2016.

France is probably the country who has extended more over the years the scope of junior-kindergarten, since 2019 all 3-year-olds should be enrolled in a school. Moreover, a large proportion of students have parents with low levels of education and/or socioeconomic status. On average, students have low z-scores, even though children do attend intensively preschool childcare. Very few estimates are significant, except for treatments  $T_{HM}$  (0.08-0.08 SD) in 2006,  $T_{ML}$  (0.13 SD) and  $T_{HL}$  (0.22-0.20 SD) in 2016.

### 6.1.3 Ontario, Québec

The two Canadian provinces have trends similar to those in Finland relatively to preschool childcare attendance, and z-scores comparable with Nordic countries, except Norway's 4<sup>th</sup> graders. Out of the nine jurisdictions considered, Ontario is the region with the highest proportion of children not attending preschool or with the lowest duration of preschool education. But the only Canadian province with a half-day junior-kindergarten –for children aged 4 years - in a public school setting. Estimates of treatment effects are generally insignificant except for  $T_{ML}$  (0-10-0.12 SD) and  $T_{HL}$  (0.12-0.13 SD) in 2011, and with the 2011-2016 pooled data.

Québec data is often used in the research literature because of a public program introduced gradually in 1998 with a very low full-time and full-year childcare fee for 0-4-year-olds not in kindergarten. Childcare enrollment surged with concomitantly large increases in mothers' labor force participation, but with rather negative impacts on children development indicators and parental well-being. Significant positive treatment effects are estimated for  $T_{HM}$  (0.17 SD) and  $T_{HL}$  (0.10-0.14 SD) with 2011 and pooled data.

In summary, treatments effects are more significant in the latest surveys, in particular the higher duration compared to the low and middle ones. Moreover most of the time there is a sorting 'gain' when comparing positive significant effect between predicted and observed score for a specific level.

### 6.2 Math and science ATETs 2015, and 2019

With the TIMSS surveys, the estimations can evaluate the impact of preschool duration on the math and science z-scores. Results are presented in the second and third panels of Table 2 for participating jurisdictions in 2015 and 2019. Again a large number of treatment effect parameters are estimated. It is not surprising that many of them are not significant if the main influence to acquire an open mind on numeracy and early awakening to science topics comes from parental preschool activities (a control in the estimations) rather than attendance to preschool childcare. Nevertheless, there are large significant effects (0.20-0.40 SD) of attending a higher rather than a lower level of preschool duration ( $T_{HL}$ ) are noticeable in math and science for Denmark, Flemish Belgium, and France and science for Sweden. In addition, a high instead of a medium level of preschool attendance ( $T_{HM}$ ) results in increases of 0.15-0.25 SD in z-scores for students in science in Denmark, and math and science in Flemish Belgium. It is the latest survey (and pooled surveys) that conduct to rather larger effects that for the reading achievement. This may also reflect that more intensive preschool activities offer very different exposition than those associated with the reading achievements the main parental domain. As for results on reading for Finland, longer enrollment ( $T_{ML}$  and  $T_{HL}$ ) leads to negative significant treatment effects. The disparities of estimates among regions may reveal that preschool is not an organization that prepares the children well to these subjects, and that schools perform differently among regions on math and science subjects.

## 7. Sensibility of ATETs to alternative durations of preschool 2015, 2016, and 2019

The categories of preschool duration proposed in the parental questionnaire by the surveys have changed over time, with the addition of a 4-year preschool enrollment for the 2015 and 2019 (TIMSS) and the 2016 (PIRLS) surveys. These changes reflect policies expansion of public policy in favor of formal childcare. They are also in phases with social behaviors related to parental socioeconomic status documented in Tables A1-A6 over the years. Estimates were conducted with two other segmentation intensities from the parental responses on preschool duration in years to analyze the sensibility of treatment effects to alternative specifications of preschool durations. In particular, the small number of children in the low duration category for many regions may have significantly impacted the estimates presented in the preceding section by increasing the standard errors of the estimated coefficients. The three specifications of the preschool duration variables are D1 (results of Table 2), and in Table 3 with the results for D2 and D3 by survey year are:

Table 2 D1: (2006-2016 2015-2019)	(1) $0 \leq D \leq 1$ ; (2) $1 < D < 3$ ; (3) $D \geq 3$ ;
Table 3 D2: (2015-2016-2019):	(1) $0 \leq D \leq 2$ ; (2) $D = 3$ ; (3) $D \geq 4$ ;
Table 3 D3: (2015-2016-2019):	(1) $0 \leq D \leq 1$ ; (2) $1 < D \leq 3$ ; (3) $D \geq 4$ .

In the D2 specification, the number of children increases in the lower category of preschool years (with a 2 years or less duration) and is reduced in the higher duration category now 4 years or more; while in the D3 specification, the number of children is kept unchanged in the first category (duration 1 year or less), is raised with a duration more than 1 year and less or equal to 3 years, and the number is reduced in the last one with 4 years or more. Estimated treatments effects are presented in Table 3 for PIRLS and TIMSS surveys.

For reading estimates, Finland is again an exception with significant negative treatment coefficients but only for the D3 segmentation which is very similar to the D1. For most jurisdictions, except for Ontario for all treatments and both segmentations (D2 and D3), and Québec with significant coefficients only for treatment T3|T1 with D2 and D3, treatment coefficients are positive and significant, many times for both specified durations (D2 and D3). There are some differences of significance between jurisdictions relatively to T2, T3, and T3 versus T1 and T2. For France and Flemish Belgium, the estimates generally indicate strong effects for all treatments (T2|T1, T3|T1, and T3|T2) and both enrollment segmentation (D2 and D3), except for one Flemish Belgium estimate (D3 and T2|T1).

For reading estimates, as seen by comparing Tables 2 with Table 3, the changes to the specification of preschool attendance durations generally result in a reduction, often significant, of standard deviations of estimated coefficients and thus, an increase in the accuracy of estimated treatment effects. A high level of preschool attendance rather than a low level ( $T_{HL}$ ) results in increases in z-scores which are relatively large (0.25-0.35 SD) for students in Denmark and France, and lower (0.10-0.20 SD), but significant, for students in Sweden, Flemish Belgium and Québec. For students in Denmark, Norway, Flemish Belgium and France, a high instead of

medium level of preschool attendance ( $T_{HM}$ ) is associated with increases of 0.08-0.15 SD in z-scores. Finally, a medium level of preschool duration rather than a lower one ( $T_{ML}$ ) results in increases of 0.12-0.20 SD in z-scores in Sweden, Flemish Belgium and France. As in Table 2, for Finnish students, treatment effects  $T_{ML}$  and  $T_{HL}$  are negative and significant for specification D3 only.

The treatment effect estimates for math and science, with specifications duration D2 and D3 presented in Table 3, lead to similar results by region. Significant effects are generated for Denmark, Flemish Belgium, and France: 0.15-0.30 SD for  $T_{HL}$  and 0.08-0.15 SD for  $T_{HM}$ . In addition, treatment effects  $T_{HM}$  become significant at that same level for math and science in Norway and math in Québec. Finland has again many negative significant coefficients associated with treatments  $T_{ML}$  and  $T_{HL}$ , and positive but not significant effects for treatment  $T_{HM}$ , for both math and science. The jurisdictions who have implemented more structured childcare systems (Denmark, Belgium, France, and Québec) present more associated positive effects.

In summary, the D2 and D3 duration specifications reflecting changes introduced in the surveys since 2015 may convey better the experience and choices of parent for their child preschool participation.

## 8. Average treatment effect by program types

For the PIRLS 2015 and TIMSS 2016 and 2019 surveys, the preschool enrollment questions were changed significantly, introducing explicitly age at enrollment, and implicitly a distinction between formal childcare and preschool. Two choices were offered in the questionnaire with a reformulated question on duration: *Did your child attend the following before first grade?*

- A. Early childhood educational program or center for children under age 3 (Yes/No);
- B. Pre-primary educational program for children aged 3 or older, including kindergarten (Yes/No);
- C. Approximately how long (in years) was your child in these programs altogether?

These participation variables have measurement errors, because the “educational” qualification in the question may have confused parents relatively to their past childcare choices. Many parents declared that their child did not participate in any program or did not know if she did, but also responded nonetheless that their child attended preschool within the duration range of less than 1 year to 4 years or more. Also, some parents declared that their child participated in only one program and responded for the duration question “more than 3 or 4 years”, which is inconsistent unless the child was cared for at home. The frequencies of children who attended two programs (for those aged less than 3 years and for the 3-year-olds or more), observed in the TIMSS 2015, 2015, and PIRLS 2016 surveys are rather high in all regions as show in the last panels of Tables D1 and D2 in section 3.

The new information introduced in the latest surveys allows for the construction of preschool program participation categories by type, and to analyze another dimension of heterogeneity. Following a four categories classification, the analysis of participation in preschool to one or two



educational programs is geared to the age of children rather than duration. Thus, a second treatment variable was created with four categories:

1. Child did not participate to any program (T1=No).
2. Child participated only in an early age childcare program (T2=Early) (could be during 4 years or more).
3. Child participated only in a late program (T3=Late).
4. Child participated to both types of programs (T4=Both).

Missing responses for questions on the types of program are higher than for checkboxes on duration, with some incoherence in the answers: responses with some (or many) years of attendance, and “child did not participate in the program” identified, or missing participation responses. The formulation of the question, as “enrolled in an early childhood educational program or center for children” may have confused some parents. A large majority of children participated in both programs. See statistics in last panel of Tables D1-D2.

Using the same control variables (described in section 3) as in section 6, only average treatment effects (ATE) are estimated to characterize the two types of programs, and to qualify the impact of using both programs compared to one and with overall years of enrollment in all programs. Table 4 presents treatment estimation results by region for reading (2016) and math/science (2015 and 2019) respectively, with pooling estimates for 2015-2019. The ATEs in Table 4 presents six ATE estimates to compare the effects of both programs, combination of both, and no participation in any program: T2early|T1no, T3late|T1no, T4both|T1no, T3late|T2early, T4both|T2early, and T4both|T3late.

Expecting that participation in both preschool programs would be higher compared to no participation, and that participation in a program for older children, a new added category, which may be more oriented toward educational development, the following ordering could potentially be observed in the estimates:

- Ranking 1: Both program/No (T4|T1) > Late program/No (T3|T1) > Early program/No (T2|T1);  
Ranking 2: Both program/Late program (T4|T3) > Both program/Early program (T4|T2) >  
Late program/Early program (T3|T2)

Table 4 shows<sup>16</sup> such rankings with significant effects, but the estimates are not systematic across regions. France, Ontario and Québec indicate positive, rather increasing, responses relatively to no participation in preschool programs. Norway, Sweden, France, and Flemish Belgium have significant treatment effects (Both/Early, Both/Late) showing that being enrolled in two programs is associated with higher z-scores rather than having been registered on one of the other program. One estimate for Ontario is interesting because early kindergarten (age 4) and kindergarten (age 5) were offered in public school till 2015-2016 (but only for half-day during school year and full-time since). France shows similar significant results for long junior

---

<sup>16</sup> For Denmark, estimates were not convergent because of the very low percentage of children in the two categories no and early program, while numbers are very high for the category both programs.

kindergarten (Late/No, Both/No), and so does Québec, the only province along with New-Brunswick and Nova Scotia offering full-time kindergarten since 1998. One shortcoming of the estimates is the blurry response on programs participation by parents.

Results for TIMSS 2015, 2019 and the pooling of years on programs participation ATEs effects presented in Table 4 indicate that treatment effects which are positive and significant are more generalized for all jurisdictions including, with the exception of Ontario and Québec. Notably, the effects  $T4|T1 > T3|T1$ ,  $T4|T2 > T3|T2$ , and the large  $T4|2$  impacts, seem to capture more the influence of European programs, as they seem to better function than the results presented relatively to reading z-scores.

## 9. Difference in treatment effects by parental education

Why do estimates of duration treatment effects present such a fragmented picture (e.g. changes of sign from one survey to another) even when the data from latest survey or pooled years characterize preschool duration? One technical explanation is to invoke that the hypothesis of confoundedness is not satisfied. This is difficult to prove considered the large diversity of jurisdictions and data sets. Another avenue is to explore if treatment effects differ with the parental education levels that is they are heterogeneous in function of education. Table 5 presents average multivalued treatments effects of treated (ATET) estimates of three parental education's levels on reading z-scores, by jurisdiction and years pooling (2011 and 2016), using a WLS-DR-ordered logit.<sup>17</sup>

Treatment effects change with parental education levels. In general they are significant only for children belonging to only one education level. For Finland, children whose parental education is high, the  $T_{ML}$  and  $T_{HL}$  are negative and significant. For Denmark, the effects  $T_{ML}$  and  $T_{HL}$  are positive and significant only for children whose parental education is low. For Ontario the effects  $T_{ML}$  and  $T_{HL}$  are positive and significant only for children of parents with a low education level. For Québec, for children of parents with the lower education level, the effects  $T_{ML}$  are positive and marginally significant; while those with parents in the higher education level have treatment effects  $T_{ML}$  et  $T_{HL}$  which are positive and significant.

Over the years, in every region selected, preschool enrollment has increased, and so has duration. Results indicate heterogeneity in intention-to-treat effects caused by the differential take-up of children and heterogeneous responses to childcare attendance.<sup>18</sup> The differences in childcare attendance by duration in terms of parental education levels are presented in Table A1.5 for each survey, except TIMSS 2011. We note that for PIRLS 2006, in several jurisdictions, the largest proportion of children with low levels of childcare attendance came from a privileged

---

<sup>17</sup> Estimations were also conducted for the math and science for 2015-2019. They show similar trends but reported less pronounced changes related to education.

<sup>18</sup> Abstracting from very recent research with data-driven methods proposals to estimate treatment effects for subpopulations, the sparse research on heterogeneity in returns to childcare typically focuses on treatment heterogeneity in observed characteristics, or estimates quantile treatment effects (QTE) rather than marginal treatment effects.

family environment with highly educated parents instead of a disadvantaged family with a low level of parental education but more potential benefits from attending quality childcare. With the exceptions of Belgium and France, children with a high level of preschool duration came predominantly from families with highly educated parents. In PIRLS 2016, children in the higher duration category have parents with a high education level while children in the lower duration category have a higher proportion, compared with 2006, of poorly educated parents. However, when comparing years 2006 and 2016, the number of children in the low preschool group has decreased significantly. These trends reflect the changing childcare environment in the selected countries. They may account for some negative treatment effects (M versus L, H versus L) in 2006 and positive but insignificant treatment impacts (H versus L, H versus M) in later years. The interpretation of the estimated treatment effects over time is made more complex by the fact that the quality of childcare services received by children has changed regardless of their duration, that the parents' education level has increased and the number of children with higher resistance to early childcare enrollment coming from a disadvantaged family has decreased.

#### **10. Social status gradients versus preschool duration gradient**

A socioeconomic gradient describes the relationship between a social outcome and the social status of individuals in a specific community. Usually, simple socioeconomic gradients are presented with three components: their level, slope, and the strength of the outcome-status relationship. In two Appendix of Figures (I for reading z-scores in PIRLS 2016 survey; II the other for math z-scores in the TIMSS 2019), for each jurisdiction and two outcomes reading or math<sup>19</sup>, four simple socioeconomic gradients slopes, estimated by an OLS estimation, are illustrated. Levels and slopes are depicted in a straightforward manner in the Figures, where the z-scores are associated three variables measuring social status and the preschool duration in years, where all variables are standardized with a mean of 0 and a standard deviation of 1.

Social status is defined by (see Tables B and C): 1. Number of parental education years (8 categories from 6 to 21). 2. Parental occupational socioeconomic index (11 categories from 22 to 73). 3. Home learning resources scale (from 5.7 to 14.8) combines data reported by students and their parents. This last scale is computed by the surveys statisticians who combine parents education years and occupations, number of children and parents books amount in home (5 categories), and computer-tablet with internet connection. Preschool duration is coded in 6 categories from 0 year to 4 or more years, for non-missing parental responses. The OLS estimates of the z-scores are conducted separately for each of the four estimators, with very few strictly exogenous covariates: students' sex, age in months, and home language if different from the test (or immigration status if available). All Figures present at the bottom the estimated coefficients, their standard errors and t-statistics. For estimations, standard errors are clustered by students' school identification, controlling for the survey sampling design of schools.

---

<sup>19</sup> To save space, Figures for science are too similar to math to be presented.

Estimated as well as observed standardized z-scores are 0 on average. Estimated gradients slopes are statistically significant at around 0.15-0.35 ( $p < 0.01$ ). Thus, students (with a mean value of the z-score) have their score increased by about 0.20-0.30 standard deviation for each increase of one standard deviation in the status index. In general, the parental years of education gradients are higher for most jurisdictions, while the occupational gradients are less steep. The home learning resources scale – likely a nearer marker of family income – is the steepest gradient. The duration gradients are much flatter than the three others and are flat in many cases. The z-scores for math (science is not presented here being too similar) have generally the same shape as for reading z-scores.

## 11. Discussion, weaknesses, and policy implications

It is well documented that students' socio-economic background is a strong predictor for learning outcomes. However, there is more uncertainty about how these background factors should be measured and how different types of measures and analytical methods influence the results. Another discussion related to social background is how comparable countries' (or jurisdiction in the same country, e.g. Ontario and Québec in Canada, Flemish and French Belgium) differences can be explained. A further important and necessary topic, arising from large-scale outcomes assessments of education, is how to improve education to achieve important educational goals, and to secure equity and inclusion for all. The analysis navigated through these challenging issues exploiting the unique rich and simultaneous information made available by the PIRLS and TIMSS multiyear surveys on Grade 4 students. These jurisdictions - countries and provinces - are diverse culturally, by their history, their policies, their economic environment, their parental background (education and occupation), with fairly desegregated enrollment of preschool or parental preschool involvement in literacy and numeracy activities. But, they are characterized with rather similar economic environment, pro-family policies, and valued education systems.

Our findings based on modelling treatment as an ordered choice of attending childcare - for either 1 year or less, between 1 and 3 years or 3 years and more rather than as a binary decision (e.g. home-care versus center-care) - does not neatly reconcile the seemingly contradictory results of positive effects for programs targeting disadvantaged children but more mixed effects for pre-school universal programs. Sooner and longer preschool educational attendance (as measured by number of years), does not appear to clearly yield strong individual and societal benefits (measured by test scores and score gaps) after children are observed at ages 9-10 in Grade 4.

Some explanations can be given regarding the mixed results on length of preschool attendance and its returns to skills. One relates to the framework used. To calculate impacts, early and more traditional childhood education research often compares a group of children who receives the intervention of interest to a group of children who receives home care ('no treatment') or in different care settings [see Casio 2021]. But often, non-random selection characterizes the enrollment of persons in a program which must be controlled, which could

raise heterogeneity in gains depending on treatment decision. The multivalued framework allows estimating the distribution of treatment effects in the population more fully. In this paper, estimated differential impacts of early childhood intervention before Grade 1 for various durations of preschool attendance were supported by data from multiple PIRLS and TIMSS large scale international surveys on test scores in Grade 4.

Over the years, in every jurisdiction selected, enrollment has increased, and so has duration. Results indicate heterogeneity in intention-to-treat effects caused by the differential take-up of children (partly linked to parental education), and also by heterogeneous responses to child care attendance.<sup>20</sup> These differences in childcare take-up and simultaneously in the impacts of uptake can be illustrated by the statistics in Table A1.5 where percentage distributions of parental education are tabulated by preschool years, for survey years. In 2006, for a few jurisdictions (e.g. Denmark) more high education parents are potentially assigned (by the model) in the low preschool category than the low education parents group, the latter usually form a more disadvantaged family background with potentially with more benefits if attending quality childcare. From 2006 to 2016, in all jurisdictions, we generally observe a large decrease in the proportion of children with a low preschool duration whose parents have a low education status. And, as expected, children assigned in the high preschool duration category, with a few exceptions, have predominantly highly educated parents. The constant general increase in education of families with younger children probably mirrors this.

These trends reflect the changing childcare environment in the selected jurisdictions. They may account for some of the negative ATETs treatments (2 versus 1; and 3 versus 1) ATETs in 2006; and the positive but not significant ATETs (3 versus 2 or 3 versus 1) treatment impacts in later years. The control groups (low and medium childcare intensities) may have shifted many students between different sorts of preschools without altering their exposure to preschool services. While at the same time, the educational level of parents has increased (see descriptive statistics in the Appendix Tables), and the number children with higher resistance to early child care enrollment coming from a disadvantaged family background have decreased.

Other aspects of the analysis could have weakened and diluted the treatment effects. Depending on year and jurisdiction, a varying number of parents did not reply to the questionnaire, which generated blank information on vital data (education, occupations, ECE enrollment and duration) which could be not imputed.<sup>21</sup> Un-confoundedness (omitted non-observable variables affecting both probabilities of treatments and score tests) could be at stake here. In particular, observed parental education and SES status variables may not capture completely parental decisions of preschool enrollment duration. Controlling for the selection into

---

<sup>20</sup> Abstracting from very recent research with data-driven methods proposals to estimate treatment effects for subpopulations, the sparse research on heterogeneity in returns to childcare typically focuses on treatment heterogeneity in observed characteristics, or estimates quantile treatment effects (QTE) rather than marginal treatment effects.

<sup>21</sup> An imputation exercise relying on chained equations to predict some less strategic missing parental variables did not seem to affect results significantly.

childcare may be crucial. Estimations that do take these selection issues into account may produce completely opposite results compared to estimations that do not (as illustrated by studies, mentioned in the literature survey, based on marginal treatment effect recent estimates in Germany and Norway). Nonetheless the paper's estimation model takes treatment duration into account with controls. But expansion of ECE policies over years, such as fees, spaces offering, and their perceived quality by parent, weight on selection decisions,<sup>22</sup> and those influence estimated treatment effect parameters. Parents must apply for care, and must be able to secure a space. Childcare research identify that a rationing mechanism is in place in almost every countries. Variations in preschool (and school) composition, quality peers and K-4 teachers, three factors which would tend to reinforce pre-existing inequalities in children's abilities, could not be controlled for in this study. In fact, there is evidence that children from disadvantaged backgrounds attend lower quality childcare and schools.

Beside extensions (see conclusion section) to tackle some of the weaknesses just identified, results can motivate two policies to generate greater equality of opportunity, that is, literacy-numeracy abilities before/at school onsets, and in early grades. Firstly, programs dedicated to enhance "good parenting behaviors" have been shown to improve the long-term success of disadvantaged children by promoting healthy maternal behaviors and by fostering parenting skills. For example, Nurse-Family Partnership programs provide home visits with nurses to disadvantaged, first-time mothers, from pregnancy until two years after birth. Secondly, quality and efficiency of teachers, principals and schools, is the other pathway to effective literacy, numeracy, social skills, and school graduation.

## 12. Conclusion

Introduction and expansions of universal ECE programs, rather than targeted ones, have been recommended to level the playing field of skills achievement and school performance in early grades among young children from disadvantaged families. On the other hand, many advantaged families (see section 4 for statistical analysis), commend such programs, having high preference for universal childcare, generally with good quality, conducive to good developmental effects, and facilitating access to the labor market with long operating hours. Besides the offering of safe and low fees services, expansion enables to secure more easily a space, escaping whatever rationing mechanism in place.

This paper has investigated the pattern of selection into the duration of childcare attendance as well as its treatment effects on international scaled test scores at Grade 4, for many countries and two Canadian provinces, having such programs and rather large extensive family policies. General expansions of programs were observed from 2006 to 2019, and significant increases of participation and enrollment duration for the succeeding cohorts of preschool children. In most

---

<sup>22</sup> Decisions could be dictated by innate abilities or sex of children. Preliminary analysis by sex suggested not adopting this distinction.

of the selected sample of jurisdictions, participation into early childcare and number of years of attendance before Grade 1 is high and has increased in the last decade.

Using the multivalued treatment effects framework, results indicate positive average treatment effects on z-scores mostly for  $ATET_{HM}$  and  $ATET_{HL}$  with much heterogeneity across jurisdictions and years. Some effects are relatively high at around 0.2 to 0.3 standard deviations, although many are not statistically significant. Some  $ATET_{ML}$  and  $ATET_{HL}$  are significantly negative, (in Finland and, marginally for Norway, and Denmark in PIRLS 2006) signaling resistance to treatment.

There is evidence of positive selection on observable gains coming from the fact that children with higher parental education are more likely to be selected into treatments with higher duration. That said, children with a less advantaged status coming from their parental education or occupation are enrolled in childcare for a smaller number of years and more likely after age 3. This evidence is more accentuated when the durations are transformed into preschool program type participation by ages.

Perhaps two lessons can be learned from this multi-country analysis. First, observable cognitive skills achievement near the end of primary school remain significant by socioeconomic gradients, in spite of major family policies to support children. Parental education proves to be a strong and largely inalterable channel through which school success is reproduced across generations. Second, the assertion that universal childcare reduces young children initial differences based on observable parental endowments, that is, their equalizing potential to lower social inequalities in educational achievement, is somewhat overstated from a policy perspective. Children of more educated parents have relatively less to gain from preschool attendance than the ones of low-educated parents, although they are more likely to be selected by parents for early childcare. Benefits may also be lower for children who have more involved parents, while parental involvement in pedagogical activities (see above), stimulating children's abilities, are codetermined by parental education. There is a strong positive interaction between the two.

Further extensions to this analysis might prove very useful to determine preschool intensity of attendance. Some extensions of the analysis would be very useful for preschool intensities to know how it goes. Because of the two-three years pandemic context with schools and ECE closures, there no new surveys. The PISA 2021 survey was postponed to 2022 with results promised end of 2023, while the older PIRLS (2021) results are coming late 2023. New education achievements linked to post-pandemic environment will take time. A new TIMSS test is scheduled for 2023, a PIRLS for year 2026, and a PISA survey in 2025.

Examining the importance of teachers' quality and their pedagogy/learning and teaching strategies, as well as students' motivation and parents' attitudes towards school and the learning process, based on the next data sets, would be important to understand their achievements in 4<sup>th</sup> Grade. Further research is needed to tap into the reasons behind differential preschool benefits across nations.

## References

- Agostinelli, F. & G. Sorrenti, G. (2021). Money vs. time: Family income, maternal labor supply, and child. University of Zurich, Department of Economics, Working Paper No. 273.
- Aizer, A. & Cunha, F. (2012). The Production of Human Capital: Endowments, Investments and Fertility, NBER Working Paper No. 18429
- Aliverninia, F. & Manganellia, S. (2015). A Multilevel Structural Equation Model Testing the Influences of Socio-Economic Status and Pre-Primary Education on Reading Literacy in Italy, *Social and Behavioral Sciences* 205: 168 – 172.
- Andresen, M. (2019). Child care for all? Treatment effects on test scores under essential heterogeneity, Equality, Social Organization, and Performance (ESOP), Statistics Norway.
- Andresen, M. (2018). Exploring Marginal Treatment Effects: Flexible estimation using Stata, *Statistics Norway, Stata Journal*, 18(1): 118-158.
- Apps, P., Mendolia, S. & Walker, I. (2013). The impact of pre-school on adolescents' outcomes: Evidence from a recent English cohort, *Economics of Education Review* 37: 183–199.
- Araujo, L. & Costa, P. (2015). Home book reading and reading achievement in EU countries: the Progress in International Reading Literacy Study 2011 (PIRLS), 21(5–6): 422–438.
- Ammermueller, A. (2013). Institutional Features of Schooling Systems and Educational Inequality: Cross-country Evidence from PIRLS and PISA, *German Economic Review* 14(2): 190–213.
- Arteaga, I., Humpagee, S., Reynolds, A. & Templebd, J. (2014). One year of preschool or two: Is it important for adult outcomes? *Economics of Education Review*, 40, 221-237.
- Baker, M., Gruber, J. & Milligan, K. (2019). The Long-Run Impacts of a Universal Child Care Program, *American Economic Journal: Economic Policy* 11(3): 1–26.
- Baker, M., Gruber, J. & Milligan, K. (2008). Universal Child Care, Maternal Labor Supply, and Family Well-Being, *Journal of Political Economy*, 116(4): 709-745.
- Bastian, J. & Lochner, L (2020). "The EITC and Maternal Time Use: More Time Working and Less Time with Kids?" NBER Working Paper 27717.
- Bauchmüller, R., Gørtz, M. & Rasmussen, A.W. (2014). Long-Run Benefits from Universal High-Quality Pre-Schooling. *Early Childhood Research Quarterly*, 29: 457-470.
- Berlinski, S., Galiani, S. & Gertler, P. (2009). The Effect of Pre-primary Education on Primary School Performance, *Journal of Public Economics* 93(1–2): 219–234.
- Bernal, R. & Keane, M. (2011). Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers. *Journal of Labor Economics* 29 (3): 459-512.
- Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4): 1117-1209.
- Björklund, A. & Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection, *The Review of Economics and Statistics*, 69(1): 42–49.
- Black, S., Devereux, P., Løken, K. & Kjell, S. (2014). Care or Cash? The Effect of Child Care Subsidies on Student Performance, *Review of Economics and Statistics*, 96(5): 824- 837.
- Blanden, J., Del Bono, E., Hansen, K., McNally, S., & Rabe, B. (2014). Early interventions and children's educational attainment. Evaluating the impact of free part-time pre-school education for 3 year olds in England, University of Surrey and CEP, London.
- Bradbury, B., Corak, M., Waldfogel, J. & Washbrook, E. (2015). Too Many Children Left Behind: The U.S. Achievement Gap in Comparative Perspective, Russell Sage Foundation, New York.
- Bradbury, B., Corak, M., Waldfogel, J. & Washbrook, E. (2012). In *Inequality in Early Childhood*, in J. Ermisch, M. Jantti and T. Smeeding (eds), Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage, Russell Sage Foundation, New York, 87-119.
- Brooks-Gunn, J., Markman-Pithers, L., & Rouse, C. (2016). Starting Early: Introducing the Issue, *Future of Children*, 26, 2: 3-17.



- Busse, A. & Gathmann, C. (2018). Free Daycare and its Effects on Children and their Families, IZA Discussion Papers 11269, Institute of Labor Economics.
- Carneiro, P. & Ginja, R. (2014). Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start, *American Economic Journal: Economic Policy* 6(4): 135–73.
- Carta, F. & Rizzica, L. (2018). Early kindergarten, maternal labor supply and children's outcomes: Evidence from Italy, *Journal of Public Economics*, 158: 79–102.
- Cascio, E. (2021). Early Childhood Education in the United States: What, When, Where, Who, How, and Why. NBER Working Paper 28722, April 2021. In preparation for *Routledge Handbook of the Economics of Education*.
- Cascio, E. (2020). Does Universal Preschool Hit the Target? Program Access and Preschool Impacts, *Journal of Human Resources*, published online before print, January 11 2021.
- Cascio, E. (2015). The Promises and Pitfalls of Universal Early Education, IZA World of Labor 116.
- Cascio, E. (2012). Knowledge, Tests, and Fadeout in Educational Interventions (with Doug Staiger). Revise and resubmit, *American Economic Review*. [NBER Working Paper 18038, May 2012]
- Cascio, E. (2009). Maternal labor supply and the introduction of kindergartens into American public schools, *Journal of Human Resources*, 44 (1): 140-170.
- Cascio, E. & Whitmore Schanzenbach, D. (2016). First in the Class? Age and the Education Production Function, *Education Finance and Policy* 11(3): 225-250.
- Cascio, E. & Whitmore Schanzenbach, D. (2013). The Impacts of Expanding Access to High-Quality Preschool Education, *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, 47(2): 127-192.
- Cebolla-Boada, H., Radl, J., Salazar, L. (2016). Preschool education as the great equalizer? A cross-country study into the sources of inequality in reading competence, *Acta Sociologica*: 1–20.
- Chor, E., Eckhoff, M. & Andresen, A. (2016). The impact of universal prekindergarten on family behavior and child outcomes, *Economics of Education Review*, 55(1): 168-181.
- Cordero, J., Cristóbal, V. & Santín, D. (2018). Causal Inference on Education Policies: A Survey of Empirical Studies Using PISA, TIMSS and PIRLS, *Journal of Economic Surveys*, 32(3): 878-915.
- Cornelissen, T., Dustmann, C., Raute, A. & Schönberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance, *Journal of Political Economy*, 126 (6), 2356-2409.
- Costa, P. & Araújo, L. (2018). Skilled Students and Effective Schools: Reading Achievement in Denmark, Sweden, and France, *Scandinavian Journal of Educational Research*, 62(6): 850-864.
- Crawford, C., Goodman, A. & Joyce, R. (2011). Explaining the socio-economic gradient in child outcomes: the inter-generational transmission of cognitive skills, *Longitudinal and Life Course Studies*, 2(1): 77-91.
- Cunha, F. & Heckman, J. (2007). The Technology of Skill Formation, *American Economic Review*, 97(2): 31-47.
- Datar, A. & Gottfried, M. (2015). School Entry Age and Children's Social-Behavioral Skills: Evidence From a National Longitudinal Study of U.S. Kindergartners, *Educational Evaluation and Policy Analysis*, 37(3): 333-353.
- Datta Gupta, N. & Simonen, M. (2010). Non-cognitive child outcomes and universal high quality child care, *Journal of Public Economics*, 94(1-2): 30-43.
- DeCicca, P. & Smith, J. (2013). The long-run impacts of early childhood education: Evidence from a failed policy experiment, *Economics of Education Review*, 36: 41-59.
- Del Bono, E., Francesconi, M., Kelly, Y. & Sacker, A. (2016). Early Maternal Time Investment and Early Child Outcomes, *The Economic Journal*, 126(596): F96–F135.
- Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start, *American Economic Journal: Applied Economics*, 1 (3): 111-34.

- Dumas, C. & Lefranc, A. (2012). Early schooling and later outcomes: Evidence from pre-school extension in France, in J. Ermisch, M. Jantti and T. Smeeding (eds), Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage, Russell Sage Foundation, New York.
- Duncan, G. & Sojourner, J. (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gap? *Journal of Human Resources* 48 (4), 945–968.
- Duncan, G. & Magnuson, K. (2013). Investing in Preschool Programs, *Journal of Economic Perspectives* 27(2): 109-32.
- Duncan, G. & Murnane, R. ed. (2011). Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances, New York: Russell Sage Foundation.
- Elango, S., García, J-L., Heckman, J., & Hojman, A (2016). Early Childhood Education, in Economics of Means-Tested Transfer Programs in the United States, Volume 2, ed. by Robert A. Moffit. Chicago, IL: University of Chicago Press.
- Elley, W. (1992). How in the World Do Students Read? IEA Study of Reading Literacy, International Association for the Evaluation of Educational Achievement.
- Ermisch, J., Jäntti, M. & Smeeding, T. (2012). Socioeconomic Gradients in Children's Outcomes, in Ermisch, M. Jantti and T. Smeeding (eds), Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage, Russell Sage Foundation, New York, chap. 2: 32-52.
- Esping-Andersen, G., Garfinkel, I., Han, W., Magnuson, K., Wagner, S., & Waldfogel, J. (2012). Child Care and School Performance in Denmark and the United States, *Child Youth Services Review*, 34(3): 576-589.
- European Commission (2014). Key Data on Early Childhood Education and Care in Europe. Eurydice and Eurostat Report. Luxembourg: Publications Office of the European Union. Council Recommendation on High Quality Early Childhood Education and Care Systems.
- Felfe, C., Nollenberger, N. & Rodriguez-Planas, N. (2015). Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development, *Journal of Population Economics*, 28(2): 393–422.
- Felfe, C. & Lalive R. (2018). Does early child care affect children's development, *Journal of Public Economics*, 15(1): 33–53.
- Feller, A., Grindal, T., Miratrix, L. & Page, L. (2016). Compared to what? Variation in the impacts of early childhood education by alternative care type, *Annals of Applied Statistics* 10(3): 1245-1285.
- Filatriau, O., Fougère, D. & Tô, M. (2013). Will sooner be better? The impact of early preschool enrollment on cognitive and non-cognitive achievement of children, Center for Research in Economics and Statistics (CREST) and Sciences Po, Paris Discussion Paper No. 9480.
- Fitzpatrick, M. , Grissmer, D. & Hasted, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment, *Economics of Education Review*, 30(2): 269-279.
- Fitzpatrick, M. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten, *Journal of Labor Economics* 28:1: 51–85.
- Fort, M., Ichino, A. & Zanella, G. (2020). The Cognitive Cost of Daycare 0-2 for Children in Advantaged Families. *Journal of Political Economy* 128: 1, 58–205.
- Fryer, R. & Levitt, S. (2004). Understanding the Black-White Test Score Gap in the First Two Years of School, *Quarterly Journal of Economics*, 119(3): 767-805.
- Garcia, E. & Weiss, E. (2015). "Early Education Gaps by Social Class and Race Start U.S. Children Out on Unequal Footing: A Summary of the Major Findings in Inequalities at the Starting Gate, Economic Policy Institute.
- García, G., Heckman, J., Duncan Ermini, L. & Prados, M.-J. (2016). The Life-cycle Benefits of an Influential Early Childhood Program, NBER Working Paper No. 22993.
- Ganzeboom, H. & Treiman, D. (2003). Three Internationally Standardised Measures for Comparative Research on Occupational Status, in Jürgen H.P. Hoffmeyer-Zlotnik & Christof Wolf (Eds.), Advances in Cross-National Comparison: A European Working Book for Demographic and SocioEconomic Variables, New York: Kluwer Academic Press, 159-193.

- Gill, L., Pennoni, F., Ramichini, C. & Romeo, I. (2015). Exploiting TIMSS and PIRLS combined data: multivariate multilevel modelling of student achievement, *The Annals of Applied Statistics* 2016, 10(4): 2405–2426.
- Goodman, A. Gregg, P. & Washbrook, E. (2011). Children’s educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood in the UK, *Longitudinal and Life Course Studies*, 2(1): 1 – 18.
- Grissmer, F. & Hastedt, (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment, *Economics of Education Review*, 30(2): 269-279.
- Haeck, C., Lebihan, L. & Merrigan, P. (2018). Universal Child Care and Long-Term Effects on Child Well-Being: Evidence from Canada, *Journal of Human Capital* 12(1): 38-98.
- Haeck, C., Lefebvre, P. & Merrigan (2015). Canadian Evidence on Ten Years of Universal Preschool Policies: The Good and the Bad, *Labour Economics*, 36: 137-157.
- Haeck, C., Lefebvre, P. & Merrigan, P. (2014). The power of the purse: New evidence impact on the distribution of income and expenditures within the family from a Canadian natural experiment, Working Paper, <https://grch.esg.uqam.ca/>
- Havnes, T. & Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127: 100-114.
- Havnes T. & Mogstad M. (2011). No Child Left Behind: Subsidized Child Care and Children’s Long-Run Outcomes, *American Economic Journal: Economic Policy*, 3(2): 97-129.
- Heckman, J. & Li, X. (2004). Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000, *Pacific Economic Review*, 9: 155–171.
- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes, *American Economic Review*, 103(6): 2052-2086.
- Heckman, J., Hyeok Moon, S., Pinto, P., Savelyev, P. & Yavitz, A. (2010a). The Rate of Return to the High Scope Perry Preschool Program, *Journal of Public Economics* 94(1-2): 114–128.
- Heckman, J., Hyeok Moon, S., Pinto, P., Savelyev, P. & Yavitz, A. (2010b). Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the High Scope Perry Preschool Program, *Quantitative Economics* 1(1): 1–46.
- Heckman, J. & Masterov, D. (2007). The Productivity Argument for Investing in Young Children, *Review of Agricultural Economics*, *American Agricultural Economics Association*, 29(3): 446-493.
- Heckman, J. & Vytlacil, E. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects, *Proceedings of the National Academy of Sciences of the United States of America*, 96(8): 4730–4734.
- Hirano, K. & Imbens, G. (2001). Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization, *Health Services and Outcomes Research Methodology*, 2(3-4): 259–278.
- Hogrebe, N. & Strietholt, R. (2016). Does non-participation in preschool affect children’s reading achievement? International evidence from propensity score analyses, *Large-scale Assessments in Education*, 4(2): 1-22.
- Imbens, G. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics*, 86(1): 4-29.
- Imbens, G. (2000). The role of the propensity score in estimating dose-response functions, *Biometrika* 87(3): 706–710.
- Imbens, G. & Rubin, D. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction*, New York, NY: Cambridge University Press.
- Jerrim, J. & Micklewright, J. (2014). Socio-economic Gradients in Children’s Cognitive Skills: Are Cross-country Comparisons Robust to Who Reports Family Background? *European Sociological Review* 30(6): 766–81.

- Joo, Y.-S., Magnuson, K., Duncan, G., Schindler, W., Yoshikawa, H., & Ziol-Guest, K. (2020). What Works in Early Childhood Education Programs? A Meta-Analysis of Preschool Enhancement Programs, *Early Education and Development*, 31(1): 1-26.
- Jones, L., Milligan, K. & Stabile, M. (2018). Child Cash Benefits and Family Expenditures: Evidence from the National Child Benefit, *Canadian Journal of Economics*, 52: 1433-1463.
- Kalil, A. & Ryan, R. (2020). Parenting Practices and Socioeconomic Gaps in Childhood Outcomes, *Future of Children*, 30(1): 29–54.
- Kalil, A., Ryan, R. & Corey, M. (2012). Diverging Destinies: Maternal education and the developmental gradient in time with children. *Demography*, 49 (4): 1361–1383.
- Kline, P. & Walters, C. (2016). Evaluating public programs with close substitutes: The case of Head Start, *Quarterly Journal of Economics*, 131(4): 1795–1848.
- Kottelenberg, M. & Lehrer, S. (2017). Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Childcare, *Journal of Labor Economics*, 35(3): 609-653.
- Kottelenberg, M., & Lehrer, S. (2013). New evidence on the impacts of access to and attending universal child-care in Canada, *Canadian Public Policy*, 39(2): 263-286.
- Ladd, H., Muschkin, C. & Dodge, K. (2014). From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina, *Journal of Policy Analysis and Management*, 33(1): 162-187.
- Lebihan, L., Haeck, C., Lefebvre, P. & Merrigan, P. (2022) "Long-term impact of a universal child care policy on parental health, behaviors and parenting practices, *The B.E. Journal of Economic Analysis and Policy*, 22(2): 311-360.
- Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption, in *Econometric Evaluation of Labour Market Policies*, Lechner M., Pfeiffer F. (eds). Physica: Heidelberg, 43–58.
- Lefebvre, P. & Merrigan, P. (2008). Childcare Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada, *Journal of Labor Economics*, 26(3): 519-548.
- Lefebvre, P., Merrigan, P. & Verstraete, M. (2009). Dynamic Labour Supply Effects of Childcare Subsidies: Evidence from a Canadian Natural Experiment on Universal Child Care, *Labour Economics*, 16(5): 490-502.
- Leuvena, E., Lindahlb, M., Oosterbeekc, H. & Webbinkd, D. (2010). Expanding schooling opportunities for 4-year-olds, *Economics of Education Review* 29: 319–328.
- Li, W., Duncan, G., Magnuson, Holly, K., Schindler, S., Yoshikawa, H., & Leak, J. (2020). Timing in Early Childhood Education: How Cognitive and Achievement Program Impacts Vary by Starting Age, Program Duration, and Time Since the End of the Program. (EdWorkingPaper: 20-201). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/5tvg-nt21>.
- Li, W. & Konstantopoulos, S. (2017). Class Size Effects on Fourth Grade Mathematics Achievement: Evidence From TIMSS 2011, *Journal of Research on Educational Effectiveness*, 28(2): 292-313.
- Linden, A., Uysal, D., Ryan, A. & Adams, J. (2016). Estimating causal effects for multivalued treatments: a comparison of approaches, *Statistics in Medicine*, 35(4): 534–552.
- Lipsey, M., Farran, D., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children’s achievement and behavior through third grade, *Early Childhood Research Quarterly*, 45(4): 155-176.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and Procedures in PIRLS 2016*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 Assessment Framework (2nd ed.)*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/framework.html>
- Mullis, I., Martin, M., Foy, P. & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

- Mullis, I., Kennedy, A., Martin, M. & Sainsbury, M. (2004). PIRLS 2006 Assessment Framework and Specifications, International Association for the Evaluation of Educational Achievement.
- Pelletier, J. & Corter, J. (2019) A longitudinal comparison of learning outcomes in full-day and half-day kindergarten, *The Journal of Educational Research*, 112:2, 192-210.
- Phillips, D., Lipsey, M., Dodge, K., Haskins, R., Bassok, D., Burchinal, M. , & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects*, Washington, DC: Brookings Institution.
- Phillips, M. (2011). Parenting, Time Use, and Disparities in Academic Outcomes, in Duncan, G. & Murnane, R. ed., *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, New York: Russell Sage Foundation, chap. 10, 207-228.
- Princiotta, D. & Germino-Hausken, E. (2006). Findings from the Fifth-grade Follow-up of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics
- Putnam, R. (2015). *Our Kids: The American Dream in Crisis*, Simon & Schuster.
- Rasmusson, M. (2016). A multilevel analysis of Swedish and Norwegian students' overall and digital reading performance with a focus on equity aspects of education, *Large-scale assessments in education*, 4(3): 2-25.
- Reeves, R. (2017). *Dream Hoarders: How the American Upper Middle Class Is Leaving Everyone Else in the Dust, Why That Is a Problem, and What to Do About It*, Brookings Institution Press, 2017.
- Reeves, R., & Howard, K. (2014). *The Parenting Gap*. Washington: Brookings Institution.
- Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70(1): 41–55.
- Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, 66: 688–701.
- Uysal, D. (2015). Doubly Robust Estimation of Causal Effects with Multivalued Treatments: An Application to the Returns to Schooling, *Journal of Applied Econometrics*, 30(5): 763-786.
- Santína, D., & Siciliab, G (2018). Does preschool education attendance matter? Evidence from a natural experiment in Spain, *Applied Economics*, 50(47): 5050–5063.
- Schütz, G. (2009). Does the quality of pre-primary education pay off in secondary school? An international comparison using PISA 2003, ifo Working Paper Series 68, ifo Institute - Leibniz Institute for Economic Research at the University of Munich.
- van Huizen, T. & Plantenga, J. (2015). *Universal Child Care and Children's Outcomes: A Meta-Analysis of Evidence from Natural Experiments*, Utrecht School of Economics Utrecht University, Utrecht, Tjalling C. Koopmans Research Institute Discussion Paper Series 15-13.
- Waldfogel, J. & Washbrook, E. (2011). Income-Related Gaps in School Readiness in the United States and the United Kingdom, in *Persistence, Privilege, and Parenting: The Comparative Study of Intergenerational Mobility*, ed. T. Smeeding, R. Erikson, and M. Jäntti , New York: Russell Sage Foundation, 175–208.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills, *Child Development* 84(6): 122–154.
- Whitehurst, G. (2018a). Does state pre-K improve children's achievement? *Early Childhood Research Quarterly*, 45(4): 183-187.
- Whitehurst, G. (2018b). The positive impacts of public pre-K fade quickly, and sometimes reverse: What does this portend for future research and policy? *Brooking Evidence Speaks Reports*, Vol 2, #59,
- Wolf, R. (1995). *The IEA reading literacy study*, Technical report, International Association for the Evaluation of Educational Achievement.

**Table 1: Percentile distribution of Grade 4 (with Grade 5 in Norway) students' overall international reading, math and science z-scores by selected jurisdiction, PIRLS 2006, 2011, 2016; TIMSS 2011, 2015, 2019**

Statistics	Reading			Math	Science	Math	Science	Math	Science
	2006	2011	2016	2011		2015		2019	
Denmark									
N	3,317	3,904	3,024	3,987	3,987	3,218	3,218	1,867	1,867
Mean	0.232	0.241	0.196	0.142	0.017	0.202	-0.002	0,138	0,067
SD	1.013	0.941	0.948	1.035	1.020	1.051	0.966	0.947	0.899
P10	-1.099	-1.017	-1.060	-1.190	-1.277	-1.160	-1.259	-1.111	-1.119
P25	-0.357	-0.344	-0.364	-0.493	-0.609	-0.481	-0.613	-0.489	-0.490
P50	0.341	0.311	0.274	0.190	0.063	0.247	0.045	0.146	0.091
P75	0.917	0.895	0.846	0.837	0.716	0.927	0.655	0.804	0.688
P90	1.430	1.379	1.330	1.416	1.265	1.519	1.181	1.321	1.182
P75-P25	1.27	1.24	1.21	1.33	1.33	1.41	1.27	1.29	1.18
Finland									
N	NA	4,463	4,506	4,223	4,223	4,704	4,704	4,146	4,146
Mean	NA	0.404	0.448	0.306	0.648	0.141	0.367	0.098	0.400
SD	NA	0.932	0.919	0.972	0.916	0.941	0.906	1.002	0.929
P10	NA	-0.813	-0.703	-0.963	-0.531	-1.093	-0.794	-1.250	-0.807
P25	NA	-0.182	-0.117	-0.322	0.077	-0.461	-0.190	-0.528	-0.171
P50	NA	0.442	0.505	0.356	0.694	0.188	0.411	0.127	0.445
P75	NA	1.038	1.074	0.967	1.266	0.794	0.974	0.793	1.036
P90	NA	1.559	1.569	1.523	1.771	1.319	1.485	1.373	1.526
P75-P25	NA	1.22	1.19	1.29	1.19	1.26	1.16	1.32	1.21
Norway Grade 4									
N	3,396	2,776	3,812	2,721	2,721	1,830	1,830	NA	NA
Mean	-0.498	-0.471	-0.236	-0.423	-0.334	-0.334	-0.318	NA	NA
SD	0.962	0.902	0.983	0.979	1.034	1.034	0.972	NA	NA
P10	-1.779	-1.650	-1.542	-1.709	-1.646	-1.646	-1.588	NA	NA
P25	-1.082	-1.057	-0.842	-1.052	-0.973	-0.973	-0.918	NA	NA
P50	-0.426	-0.414	-0.159	-0.387	-0.282	-0.282	-0.253	NA	NA
P75	0.169	0.154	0.442	0.250	0.357	0.357	0.352	NA	NA
P90	0.667	0.631	0.966	0.810	0.953	0.953	0.872	NA	NA
P75-P25	1.25	1.21	1.28	1.30	1.33	1.33	1.27	NA	NA
Norway Grade 5									
N	NA	NA	3,808	NA	NA	1,814	1,814	2,331	2,331
Mean	NA	NA	0.304	NA	NA	0.465	0.289	0.301	0.252
SD	NA	NA	0.931	NA	NA	0.990	0.882	0.991	0.937
P10	NA	NA	-0.925	NA	NA	-0.787	-0.857	-1.032	-0.967
P25	NA	NA	-0.291	NA	NA	-0.164	-0.245	-0.343	-0.319
P50	NA	NA	0.343	NA	NA	0.510	0.320	0.369	0.340
P75	NA	NA	0.928	NA	NA	1.133	0.883	1.003	0.927
P90	NA	NA	1.443	NA	NA	1.635	1.435	1.522	1.383
P75-P25	NA	NA	1.22	NA	NA	1.33	1.33	1.35	1.25

See Notes and Sources at Table 1: End.

**Table 1: Continued**

Statistics	Reading			Math	Science	Math	Science	Math	Science
	2006	2011	2016	2011		2015		2019	
	Sweden					Sweden			
N	3,468	3,646	3,582	3,578	3,578	6,622	6,622	3,053	3,053
Mean	0.277	0.093	0.339	-0.239	0.210	-0.014	0.277	0.028	0.246
SD	0.922	0.937	0.925	0.932	1.002	0.933	0.990	0.954	0.976
P10	-0.924	-1.117	-0.891	-1.452	-1.116	-1.245	-1.016	-1.211	-1.034
P25	-0.270	-0.477	-0.217	-0.840	-0.424	-0.610	-0.325	-0.578	-0.351
P50	0.326	0.136	0.414	-0.207	0.267	0.025	0.345	0.044	0.297
P75	0.900	0.710	0.972	0.394	0.894	0.628	0.948	0.687	0.910
P90	1.397	1.240	1.450	0.924	1.442	1.140	1.481	1.224	1.471
P75-P25	1.17	1.19	1.19	1.24	1.32	1.24	1.27	1.27	1.26
	Belgium Flemish					Belgium Flemish			
N	4,078	NA	4,465	4,849	4,849	4,827	4,827	4,171	4,171
Mean	0.221	NA	-0.110	0.315	-0.269	0.299	-0.210	0.093	-0.318
SD	0.801	NA	0.837	0.863	0.813	0.856	0.863	0.887	0.890
P10	-0.804	NA	-1.195	-0.808	-1.353	-0.803	-1.343	-1.056	-1.489
P25	-0.293	NA	-0.657	-0.269	-0.805	-0.274	-0.770	-0.508	-0.887
P50	0.246	NA	-0.078	0.326	-0.244	0.309	-0.175	0.108	-0.270
P75	0.760	NA	0.474	0.909	0.303	0.884	0.386	0.711	0.309
P90	1.216	NA	0.932	1.420	0.758	1.398	0.866	1.231	0.775
P75-P25	1.05	NA	1.13	1.17	1.11	1.16	1.16	1.22	1.20
	Belgium French					Belgium French			
N	3,821	3,086	3,666	NA	NA	NA	NA	NA	NA
Mean	-0.460	-0.472	-0.512	NA	NA	NA	NA	NA	NA
SD	0.995	0.947	0.961	NA	NA	NA	NA	NA	NA
P10	-1.764	-1.732	-1.765	NA	NA	NA	NA	NA	NA
P25	-1.110	-1.061	-1.132	NA	NA	NA	NA	NA	NA
P50	-0.413	-0.428	-0.472	NA	NA	NA	NA	NA	NA
P75	0.224	0.185	0.144	NA	NA	NA	NA	NA	NA
P90	0.775	0.706	0.704	NA	NA	NA	NA	NA	NA
P75-P25	1.33	1.25	1.28	NA	NA	NA	NA	NA	NA
	France					France			
N	3,763	3,927	4,061	NA	NA	4,061	4,061	3,739	3,739
Mean	-0.117	-0.274	-0.314	NA	NA	0.023	-0.529	-0.539	-0.523
SD	0.960	0.994	0.964	NA	NA	0.941	1.015	1.050	1.052
P10	-1.389	-1.586	-1.592	NA	NA	-1.197	-1.871	-1.951	-1.927
P25	-0.740	-0.925	-0.913	NA	NA	-0.579	-1.191	-1.256	-1.207
P50	-0.070	-0.213	-0.248	NA	NA	0.060	-0.482	-0.491	-0.454
P75	0.553	0.417	0.358	NA	NA	0.666	0.184	0.189	0.220
P90	1.082	0.964	0.857	NA	NA	1.202	0.733	0.788	0.767
P75-P25	1.29	1.34	1.27	NA	NA	1.25	1.37	1.44	1.43

See Notes and Sources at Table 1: End.

**Table 1: End**

Statistics	Reading			Math	Science	Math	Science	Math	Science
	2006	2011	2016	2011		2015		2019	
	Ontario					Ontario			
N	3,346	3,478	3,287	4,570	4,570	3,525	3,525	2,564	2,564
Mean	0.351	0.235	0.188	-0.142	0.016	-0.123	0.095	-0.071	0.077
SD	1.023	1.030	1.080	1.058	1.073	1.017	1.003	1.033	1.008
P10	-0.973	-1.156	-1.265	-1.501	-1.400	-1.454	-1.239	-1.435	-1.281
P25	-0.297	-0.395	-0.478	-0.818	-0.658	-0.792	-0.549	-0.810	-0.580
P50	0.387	0.300	0.286	-0.112	0.081	-0.086	0.150	-0.045	0.133
P75	1.050	0.948	0.942	0.572	0.749	0.559	0.789	0.660	0.763
P90	1.648	1.511	1.489	1.172	1.349	1.156	1.337	1.224	1.345
P75-P25	1.35	1.34	1.42	1.59	1.41	1.35	1.34	1.47	1.34
	Québec					Québec			
N	3,046	3,603	2,687	3,600	3,600	2,355	2,355	2,848	2,848
Mean	0.031	-0.008	0.193	0.121	-0.100	0.046	-0.014	0.112	-0.023
SD	0.922	0.900	0.896	0.861	0.816	0.937	0.919	0.884	0.880
P10	-1.161	-1.154	-0.961	-1.006	-1.138	-1.175	-1.181	-1.054	-1.207
P25	-0.548	-0.591	-0.392	-0.464	-0.633	-0.564	-0.615	-0.493	-0.634
P50	0.072	0.019	0.217	0.136	-0.083	0.077	-0.007	0.169	0.008
P75	0.672	0.599	0.792	0.721	0.453	0.684	0.610	0.693	0.605
P90	1.167	1.104	1.327	0.924	0.946	1.224	1.148	1.223	1.071
P75-P25	1.22	1.19	1.18	1.24	1.09	1.25	1.23	1.19	1.24

Notes: Z-scores are computed with house weights; percentile based on 5 plausible values and if student's information on parental education is not missing. Statistics: N: number of students in Grade 4 (or Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc.; P75-P25 the z-score gap between the P75 and P25 percentiles. NA: Indicates that the jurisdiction did not participate in this survey. In year 2011 when both surveys were conducted in the selected jurisdictions, some choose to have two samples of students but only one parental questionnaire. Treatments effects for math and science are not estimated in 2011 for Denmark, Belgium Flemish, and Ontario because parental covariates are missing. Shaded statistics identified these jurisdictions.

Sources: Authors' computation from weighted PIRLS and TIMSS surveys data sets.



**Table 2: Average multivalued treatments effects of the treated, ATET (std. err.) estimates of three preschool duration levels (Low, Middle, High) on Grade 4 (with Grade 5 in Norway) students reading, math and science z-scores, PIRLS 2006-2011-2016, TIMSS 2011-2015-2019, and years pooling by jurisdiction**

Scores	Year	N	ATET Middle vs Low		ATET High vs Low		ATET High vs Middle	
			T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
Denmark								
Reading	2006	3,192	<b>-0.31***</b> (0.09)	<b>-0.31***</b> (0.10)	<b>-0.24**</b> (0.09)	<b>-0.21**</b> (0.09)	<b>0.10*</b> (0.06)	0.07 (0.05)
	2011	3,721	0.21 (0.14)	<b>0.24*</b> (0.13)	<b>0.36***</b> (0.13)	<b>0.31**</b> (0.13)	<b>0.08*</b> (0.04)	<b>0.08*</b> (0.04)
	2016	2,841	0.00 (0.23)	0.03 (0.22)	0.18 (0.18)	<b>0.34*</b> (0.19)	0.28 (0.20)	<b>0.28**</b> (0.14)
	2011-16	6,562	0.19 (0.12)	0.13 (0.12)	<b>0.39***</b> (0.14)	<b>0.32***</b> (0.11)	<b>0.22***</b> (0.08)	<b>0.11***</b> (0.04)
Math	2015	2,997	0.15 (0.16)	0.11 (0.15)	0.20 (0.14)	<b>0.22*</b> (0.13)	0.18 (0.11)	0.11 (0.11)
	2019	2,641	0.05 (0.32)	-0.34 (0.28)	0.48 (0.48)	-0.06 (0.27)	0.20 (0.13)	<b>0.23**</b> (0.12)
	2015-19	5,638	-0.17 (0.17)	-0.05 (0.16)	-0.07 (0.15)	0.11 (0.13)	<b>0.21**</b> (0.09)	<b>0.19**</b> (0.08)
Science	2015	2,997	0.07 (0.15)	0.06 (0.18)	0.16 (0.11)	<b>0.19*</b> (0.11)	<b>0.17**</b> (0.07)	0.12 (0.08)
	2019	2,641	0.04 (0.29)	0.02 (0.28)	0.27 (0.69)	0.23 (0.26)	0.09 (0.11)	0.15 (0.11)
	2015-19	5,638	-0.01 (0.13)	0.05 (0.19)	0.11 (0.14)	<b>0.21*</b> (0.12)	<b>0.13*</b> (0.07)	<b>0.14**</b> (0.07)
Finland								
Reading	2011	4,346	<b>-0.11***</b> (0.04)	<b>-0.08*</b> (0.04)	<b>-0.09**</b> (0.04)	<b>-0.07*</b> (0.04)	0.02 (0.03)	0.02 (0.03)
	2016	4,363	<b>-0.08*</b> (0.05)	<b>-0.08*</b> (0.05)	<b>-0.09**</b> (0.04)	<b>-0.08**</b> (0.04)	0.01 (0.04)	0.01 (0.04)
	2011-16	8,709	<b>-0.10***</b> (0.03)	<b>-0.08**</b> (0.03)	<b>-0.09***</b> (0.03)	<b>-0.07***</b> (0.03)	0.01 (0.03)	0.02 (0.03)
Math	2011	4,105	<b>-0.13***</b> (0.04)	<b>-0.11***</b> (0.04)	<b>-0.14***</b> (0.05)	<b>-0.11**</b> (0.04)	0.00 (0.04)	0.00 (0.04)
	2015	4,529	-0.02 (0.06)	-0.01 (0.06)	-0.09 (0.06)	<b>-0.10*</b> (0.06)	-0.07 (0.05)	-0.07 (0.05)
	2019	5,669	0.00 (0.06)	0.03 (0.06)	-0.04 (0.05)	-0.04 (0.05)	-0.05 (0.05)	-0.05 (0.06)
	2015-19	10,198	0.00 (0.04)	0.01 (0.04)	-0.05 (0.04)	<b>-0.06*</b> (0.04)	-0.06 (0.04)	-0.06 (0.04)
Science	2011	4,105	<b>-0.11***</b> (0.03)	<b>-0.09**</b> (0.04)	<b>-0.12**</b> (0.05)	<b>-0.09*</b> (0.05)	-0.01 (0.04)	0.00 (0.04)
	2015	4,529	-0.06 (0.07)	-0.07 (0.06)	<b>-0.12**</b> (0.06)	<b>-0.14**</b> (0.06)	-0.07 (0.06)	-0.07 (0.06)
	2019	5,669	-0.03 (0.06)	-0.01 (0.06)	<b>-0.08*</b> (0.04)	<b>-0.08*</b> (0.05)	-0.05 (0.05)	-0.05 (0.05)
	2015-19	10,198	-0.03 (0.04)	-0.04 (0.05)	<b>-0.09**</b> (0.03)	<b>-0.10***</b> (0.04)	-0.06 (0.04)	-0.06 (0.04)

Notes: see end of Table 2.

**Table 2: Continued**

Score	Year	N	ATET Middle vs Low		ATET High vs Low		ATET High vs Middle	
			T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
Norway Grade 5								
Reading	2006	3,043	-0.02 (0.07)	-0.02 (0.08)	0.09 (0.08)	0.06 (0.07)	0.06 (0.05)	0.08 (0.06)
	2011	2,606	-0.08 (0.11)	-0.06 (0.12)	-0.01 (0.12)	-0.01 (0.14)	0.03 (0.05)	0.04 (0.05)
	2006-11	5,528	-0.02 (0.06)	0.00 (0.06)	0.05 (0.07)	0.05 (0.06)	0.05 (0.04)	0.05 (0.04)
	2016	3,582	-0.13 (0.10)	-0.15 (0.10)	0.03 (0.05)	0.01 (0.05)	<b>0.26**</b> (0.11)	0.14 (0.10)
Math	2011	2,565	<b>-0.16*</b> (0.09)	-0.15 (0.11)	-0.08 (0.13)	-0.10 (0.12)	0.01 (0.05)	0.03 (0.05)
	2015	1,736	0.26 (0.32)	0.27 (0.22)	0.17 (0.20)	0.25 (0.17)	-0.10 (0.15)	-0.09 (0.19)
	2019	3,334	0.12 (0.18)	0.23 (0.16)	0.16 (0.22)	0.18 (0.14)	0.11 (0.16)	0.02 (0.12)
	2015-19	5,129	0.18 (0.15)	<b>0.24*</b> (0.12)	0.17 (0.18)	<b>0.23**</b> (0.11)	0.04 (0.14)	0.00 (0.09)
Science	2011	2,565	-0.09 (0.08)	-0.09 (0.10)	0.01 (0.10)	-0.03 (0.10)	0.02 (0.05)	0.04 (0.05)
	2015	1,736	0.17 (0.33)	0.20 (0.24)	0.28 (0.21)	0.33 (0.20)	<b>0.18*</b> (0.10)	0.17 (0.16)
	2019	3,334	0.13 (0.18)	0.25 (0.21)	0.22 (0.20)	<b>0.35**</b> (0.14)	<b>0.25*</b> (0.14)	0.11 (0.15)
	2015-19	5,129	0.14 (0.16)	0.23 (0.15)	<b>0.27*</b> (0.15)	<b>0.37***</b> (0.11)	<b>0.26**</b> (0.11)	0.15 (0.11)
Norway Grade 4								
Reading	2006	3,124	-0.03 (0.09)	-0.02 (0.09)	0.10 (0.29)	0.08 (0.09)	0.08 (0.06)	0.10 (0.07)
	2011	2,650	-0.10 (0.14)	-0.05 (0.15)	-0.02 (0.61)	0.04 (0.17)	0.05 (0.06)	0.05 (0.06)
	2016	3,652	-0.16 (0.12)	-0.18 (0.12)	0.04 (0.43)	0.03 (0.06)	<b>0.32**</b> (0.13)	0.19 (0.12)
	2011-16	6,202	-0.10 (0.12)	-0.15 (0.11)	0.08 (0.16)	0.03 (0.03)	<b>0.21**</b> (0.08)	<b>0.19**</b> (0.07)
	2006-11-16	9,326	-0.04 (0.07)	-0.08 (0.07)	0.05 (0.05)	0.05 (0.05)	<b>0.17**</b> (0.06)	<b>0.08*</b> (0.06)
Math	2011	2,606	-0.20 (0.10)	-0.17 (0.13)	-0.10 (0.15)	-0.10 (0.14)	0.01 (0.06)	0.05 (0.0)
	2015	1,823	-0.65 (0.37)	-0.46 (0.34)	-0.16 (0.24)	-0.34 (0.23)	0.14 (0.31)	0.32 (0.25)
	2011-15	4,429	<b>-0.21*</b> (0.09)	<b>-0.23*</b> (0.12)	-0.11 (0.13)	-0.13 (0.11)	0.09 (0.08)	0.07 (0.05)
Science	2011	2,606	-0.11 (0.09)	-0.11 (0.11)	0.00 (0.11)	-0.02 (0.12)	0.05 (0.06)	0.05 (0.06)
	2015	6,288	-0.49 (0.37)	-0.33 (0.37)	-0.02 (0.25)	-0.21 (0.23)	0.32 (0.36)	0.37 (0.33)
	2011-15	4,429	-0.12 (0.09)	-0.16 (0.13)	0.03 (0.15)	-0.04 (0.11)	0.16 (0.09)	0.08 (0.06)

Notes: see end of Table 2.

**Table 2: Continued**

Score	Year	N	ATET Middle vs Low		ATET High vs Low		ATET High vs Middle	
			T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
Sweden								
Reading	2006	3,293	0.10 (0.07)	0.09 (0.08)	<b>0.17**</b> (0.08)	<b>0.17**</b> (0.07)	0.06 (0.03)	0.06 (0.04)
	2011	3,372	-0.11 (0.16)	-0.03 (0.10)	-0.10 (0.22)	0.05 (0.11)	<b>0.08**</b> (0.04)	<b>0.07*</b> (0.04)
	2006-11	6,632	0.03 (0.07)	0.07 (0.06)	0.05 (0.11)	<b>0.13**</b> (0.06)	<b>0.08***</b> (0.03)	<b>0.07**</b> (0.03)
	2016	3,442	<b>0.20*</b> (0.11)	<b>0.23**</b> (0.11)	<b>0.25**</b> (0.10)	<b>0.21**</b> (0.09)	-0.03 (0.09)	0.01 (0.09)
	2006-11-16	10,024	0.04 (0.06)	<b>0.11**</b> (0.05)	0.14 (0.08)	<b>0.16***</b> (0.05)	0.06 (0.04)	<b>0.07**</b> (0.03)
Math	2011	2,606	0.04 (0.11)	0.03 (0.08)	0.04 (0.12)	0.08 (0.08)	0.05 (0.04)	0.05 (0.04)
	2015	6,288	0.04 (0.13)	0.05 (0.15)	0.05 (0.10)	0.06 (0.11)	0.02 (0.09)	0.03 (0.10)
	2019	4,049	0.09 (0.14)	-0.01 (0.15)	<b>0.19**</b> (0.09)	-0.01 (0.14)	-0.01 (0.09)	0.01 (0.11)
	2015-19	10,337	0.06 (0.10) 7	0.00 (0.11)	<b>0.14*</b> (0.08)	0.02 (0.09)	0.00 (0.06)	0.03 (0.07)
	2011	3,300	0.05 (0.12)	0.06 (0.08)	0.05 (0.14)	0.12 (0.08)	0.06 (0.04)	0.06 (0.04)
Science	2015	6,288	<b>0.25*</b> (0.15)	<b>0.29*</b> (0.16)	0.10 (0.10)	0.16 (0.13)	-0.10 (0.10)	-0.10 (0.11)
	2019	4,049	0.21 (0.16)	0.05 (0.17)	<b>0.46***</b> (0.15)	<b>0.23*</b> (0.14)	0.12 (0.10)	0.19 (0.11)
	2015-19	10,337	<b>0.22**</b> (0.11)	0.15 (0.11)	<b>0.28***</b> (0.09)	0.16 (0.11)	0.01 (0.07)	0.01 (0.08)

Notes: see end of Table 2.

**Table 2: Continued**

Score	Year	N	ATET Middle vs Low		ATET High vs Low		ATET High vs Middle	
			T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
Belgium Flemish								
Reading	2006	3,894	0.16 (0.61)	0.22 (0.23)	0.42 (0.98)	0.19 (0.23)	<b>0.09**</b> (0.04)	0.06 (0.04)
	2016	4,166	-0.06 (0.10)	-0.09 (0.11)	<b>0.14*</b> (0.08)	<b>0.14**</b> (0.06)	<b>0.19***</b> (0.06)	<b>0.20***</b> (0.07)
	2006-16	8,060	0.05 (0.22)	-0.03 (0.10)	0.13 (0.14)	<b>0.18***</b> (0.06)	<b>0.12***</b> (0.04)	<b>0.09***</b> (0.03)
Math	2015	4,398	0.11 (0.10)	0.13 (0.10)	<b>0.19***</b> (0.07)	<b>0.21***</b> (0.06)	<b>0.12*</b> (0.07)	0.08 (0.08)
	2019	3,668	0.15 (0.10)	0.06 (0.11)	<b>0.31***</b> (0.08)	<b>0.23***</b> (0.07)	<b>0.12*</b> (0.07)	<b>0.13*</b> (0.07)
	2015-19	8,066	<b>0.13*</b> (0.07)	0.11 (0.08)	<b>0.25***</b> (0.06)	<b>0.21***</b> (0.05)	<b>0.12**</b> (0.06)	<b>0.10*</b> (0.06)
Science	2015	4,398	0.16 (0.10)	<b>0.18**</b> (0.09)	<b>0.30***</b> (0.07)	<b>0.32***</b> (0.06)	<b>0.14**</b> (0.07)	<b>0.14*</b> (0.07)
	2019	3,668	<b>0.24**</b> (0.12)	0.18 (0.12)	<b>0.42***</b> (0.11)	<b>0.30***</b> (0.08)	<b>0.13**</b> (0.05)	<b>0.13**</b> (0.06)
	2015-19	8,066	<b>0.20**</b> (0.08)	<b>0.18**</b> (0.07)	<b>0.36***</b> (0.07)	<b>0.30***</b> (0.05)	<b>0.14***</b> (0.05)	<b>0.13***</b> (0.05)
Belgium French								
Reading	2006	3,215	-0.09 (0.19)	0.13 (0.20)	-0.01 (0.19)	0.21 (0.16)	<b>0.11**</b> (0.05)	<b>0.11**</b> (0.04)
	2011	5,678	-0.09 (0.15)	-0.10 (0.16)	0.03 (0.17)	-0.01 (0.18)	<b>0.11**</b> (0.05)	<b>0.10**</b> (0.05)
	2016	3,410	0.01 (0.13)	0.14 (0.14)	0.06 (0.12)	<b>0.26**</b> (0.11)	0.03 (0.09)	0.12 (0.08)
	2011-16	9,063	-0.10 (0.16)	0.04 (0.12)	0.03 (0.12)	0.09 (0.11)	<b>0.10**</b> (0.04)	<b>0.10**</b> (0.04)
	2006-11-16	12,115	-0.12 (0.12)	0.03 (0.10)	0.01 (0.10)	0.09 (0.09)	<b>0.10***</b> (0.03)	<b>0.11***</b> (0.04)
France								
Reading	2006	3,293	0.19 (0.44)	-0.17 (0.24)	0.27 (0.56)	-0.11 (0.22)	<b>0.08**</b> (0.04)	<b>0.08*</b> (0.04)
	2011	3,372	-0.41 (0.66)	-0.21 (0.21)	-0.42 (0.60)	-0.15 (0.24)	0.01 (0.04)	0.01 (0.04)
	2006-11	6,642	-0.19 (0.29)	-0.19 (0.17)	-0.20 (0.31)	-0.12 (0.17)	0.05 (0.03)	0.05 (0.03)
	2016	3,711	0.12 (0.08)	<b>0.13*</b> (0.07)	<b>0.22***</b> (0.05)	<b>0.20***</b> (0.05)	0.08 (0.06)	0.08 (0.06)
Math	2015	3,581	-0.07 (0.11)	-0.03 (0.11)	<b>0.12*</b> (0.06)	<b>0.13**</b> (0.06)	0.15 (0.11)	<b>0.18*</b> (0.10)
	2019	4,680	0.08 (0.10)	0.07 (0.10)	<b>0.15**</b> (0.06)	<b>0.16**</b> (0.07)	0.05 (0.07)	0.07 (0.08)
	2015-19	8,261	0.02 (0.08)	0.03 (0.08)	<b>0.14***</b> (0.05)	<b>0.15***</b> (0.05)	0.09 (0.06)	<b>0.11*</b> (0.06)
Science	2015	3,581	0.03 (0.11)	0.10 (0.11)	0.11 (0.08)	0.13 (0.08)	0.05 (0.09)	0.06 (0.09)
	2019	4,680	<b>0.16*</b> (0.09)	0.15 (0.10)	<b>0.23***</b> (0.06)	<b>0.24***</b> (0.07)	0.04 (0.07)	0.07 (0.07)
	2015-19	8,261	<b>0.12*</b> (0.07)	<b>0.12*</b> (0.07)	<b>0.20***</b> (0.05)	<b>0.21***</b> (0.05)	0.05 (0.06)	0.06 (0.06)

Notes: see end of Table 2.

Table 2: End

Score	Year	N	ATET Middle vs Low		ATET High vs Low		ATET High vs Middle	
			T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
Ontario								
Reading	2006	3,053	-0.06 (0.05)	-0.08 (0.05)	-0.06 (0.07)	-0.11 (0.08)	0.00 (0.07)	-0.02 (0.07)
	2011	3,240	<b>0.10**</b> (0.05)	<b>0.12**</b> (0.05)	<b>0.12*</b> (0.07)	<b>0.15*</b> (0.08)	0.05 (0.05)	0.03 (0.06)
	2016	3,026	0.05 (0.06)	0.06 (0.06)	0.10 (0.07)	0.10 (0.06)	0.04 (0.05)	0.04 (0.05)
	2011-16	6,266	<b>0.09**</b> (0.04)	0.07 (0.05)	<b>0.10*</b> (0.05)	<b>0.12**</b> (0.05)	0.05 (0.04)	0.04 (0.04)
Math	2015	3,366	-0.07 (0.04)	-0.06 (0.04)	0.02 (0.05)	-0.03 (0.05)	0.09 (0.05)	0.08 (0.05)
	2019	2,037	0.01 (0.08)	0.02 (0.08)	0.10 (0.08)	0.06 (0.10)	0.07 (0.06)	0.07 (0.06)
	2015-19	5,403	-0.03 (0.05)	-0.03 (0.04)	0.06 (0.05)	0.01 (0.05)	<b>0.08**</b> (0.04)	<b>0.07*</b> (0.04)
Science	2015	3,366	-0.01 (0.04)	0.00 (0.05)	0.04 (0.05)	0.01 (0.05)	0.04 (0.04)	0.04 (0.04)
	2019	2,037	-0.03 (0.08)	-0.05 (0.08)	0.07 (0.06)	0.02 (0.07)	0.07 (0.07)	0.08 (0.06)
	2015-19	5,403	-0.02 (0.05)	-0.02 (0.04)	0.05 (0.04)	0.02 (0.04)	0.06 (0.03)	<b>0.06*</b> (0.04)
Québec								
Reading	2006	2,714	0.00 (0.05)	0.00 (0.04)	-0.09 (0.07)	-0.10 (0.07)	-0.09 (0.07)	-0.08 (0.07)
	2011	3,295	0.03 (0.04)	0.03 (0.04)	<b>0.17***</b> (0.05)	<b>0.17***</b> (0.06)	<b>0.14***</b> (0.05)	<b>0.14**</b> (0.05)
	2016	2,490	-0.01 (0.07)	0.00 (0.07)	<b>0.09*</b> (0.05)	<b>0.08*</b> (0.05)	0.10 (0.07)	0.07 (0.07)
	2011-16	5,776	0.03 (0.04)	0.02 (0.04)	<b>0.10**</b> (0.04)	<b>0.14***</b> (0.04)	0.09 (0.06)	<b>0.12***</b> (0.04)
Math	2011	3,294	<b>-0.08**</b> (0.04)	<b>-0.08**</b> (0.04)	-0.01 (0.04)	0.00 (0.04)	<b>0.09*</b> (0.05)	0.07 (0.05)
	2015	2,256	0.03 (0.07)	0.05 (0.07)	0.04 (0.06)	0.01 (0.06)	0.04 (0.07)	-0.03 (0.06)
	2019	1,764	-0.03 (0.07)	-0.09 (0.07)	0.10 (0.06)	0.10 (0.06)	<b>0.13**</b> (0.06)	<b>0.14**</b> (0.06)
	2015-19	4,020	0.01 (0.05)	0.02 (0.05)	<b>0.07*</b> (0.04)	0.05 (0.04)	0.06 (0.05)	0.05 (0.04)
Science	2011	3,294	-0.05 (0.03)	-0.04 (0.03)	0.04 (0.04)	0.03 (0.05)	<b>0.10*</b> (0.06)	0.08 (0.06)
	2015	2,256	0.04 (0.08)	0.05 (0.08)	0.00 (0.06)	-0.04 (0.07)	0.01 (0.08)	-0.08 (0.07)
	2019	1,764	-0.01 (0.08)	-0.05 (0.09)	0.05 (0.05)	0.05 (0.05)	0.06 (0.07)	0.06 (0.07)
	2015-19	4,020	0.02 (0.06)	0.03 (0.06)	0.02 (0.04)	-0.01 (0.05)	0.01 (0.06)	-0.02 (0.05)

Notes: ATETs are average treatment effect on the treated. Statistics: N indicates number of students in Grade 4 (Grade 5 in Norway) with test scores and parental covariates. Coefficient and (Std. Err.) indicate estimated parameters. The three treatment preschool durations used are: (Low)  $\leq 1$  year; (Medium) 1 year and  $< 3$  years; (High)  $\geq 3$  years. Statistical significance: If a p-value is less than 0.05, one star (\*); If less than 0.01, two stars (\*\*); If less than 0.001, three stars (\*\*\*)

Source: Authors' econometric model (see text).

**Table 3: Average multivalued treatments effects of the treated, ATET (std. err.) for two preschool duration's alternative levels (D2 and D3) on z-scores (PIRLS 2016), on math and science (TIMSS 2015, 2019), and years pooling by jurisdiction**

Score	D	Year	N	ATET Medium vs Low		ATET High vs Low		ATET High vs Medium	
				T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
<b>Denmark</b>									
Reading	D2	2016	2,850	0.10 (0.13)	<b>0.23*</b> (0.14)	0.24 (0.15)	<b>0.33**</b> (0.12)	<b>0.12**</b> (0.06)	<b>0.11*</b> (0.06)
	D3	2016	2,850	0.11 (0.20)	0.18 (0.20)	0.09 (0.17)	0.35* (0.19)	0.16** (0.07)	0.15** (0.06)
Math	D2	2015	2,997	0.15 (0.11)	0.09 (0.10)	<b>0.21**</b> (0.10)	<b>0.19*</b> (0.10)	<b>0.09**</b> (0.04)	<b>0.09**</b> (0.04)
		2019	2,641	0.02 (0.13)	0.03 (0.12)	0.17 (0.12)	0.17* (0.10)	0.19** (0.08)	0.15** (0.07)
		2015-19	5,638	0.13 (0.10)	0.08 (0.08)	<b>0.18**</b> (0.09)	<b>0.19***</b> (0.07)	<b>0.16***</b> (0.06)	<b>0.10***</b> (0.04)
	D3	2015	2,997	0.20 (0.14)	0.17 (0.13)	0.19 (0.14)	<b>0.26*</b> (0.13)	<b>0.10***</b> (0.04)	<b>0.10**</b> (0.04)
		2019	2,641	0.04 (0.26)	-0.14 (0.26)	0.61 (0.68)	-0.11 (0.28)	<b>0.20***</b> (0.06)	<b>0.18***</b> (0.06)
		2015-19	5,638	0.03 (0.12)	0.04 (0.13)	-0.03 (0.17)	0.13 (0.14)	0.17** (0.05)	0.12** (0.03)
Science	D2	2015	2,997	0.13* (0.07)	0.09 (0.07)	<b>0.19***</b> (0.07)	<b>0.19***</b> (0.07)	<b>0.10**</b> (0.04)	<b>0.09**</b> (0.04)
		2019	2,641	-0.04 (0.12)	-0.03 (0.12)	0.09 (0.11)	<b>0.20*</b> (0.12)	<b>0.21**</b> (0.09)	<b>0.22**</b> (0.09)
		2015-19	5,638	0.10 (0.06)	0.05 (0.07)	0.12 (0.08)	<b>0.20***</b> (0.07)	<b>0.18***</b> (0.06)	<b>0.11**</b> (0.04)
	D3	2015	2,997	0.15 (0.11)	0.13 (0.12)	0.16 (0.11)	<b>0.24**</b> (0.11)	0.11*** (0.04)	0.09** (0.04)
		2019	2,641	0.01 (0.24)	0.12 (0.26)	0.33 (0.93)	0.21 (0.30)	<b>0.16**</b> (0.07)	<b>0.20**</b> (0.08)
		2015-19	5,638	0.06 (0.10)	0.12 (0.14)	0.16 (0.15)	0.25* (0.13)	<b>0.14**</b> (0.05)	<b>0.12***</b> (0.04)
<b>Finland</b>									
Reading	D2	2016	4,366	-0.06 (0.04)	-0.07 (0.04)	-0.05 (0.03)	-0.04 (0.04)	0.01 (0.03)	0.02 (0.03)
	D3	2016	4,366	<b>-0.09**</b> (0.04)	<b>-0.09**</b> (0.04)	-0.09** (0.04)	<b>-0.07*</b> (0.04)	0.01 (0.03)	0.02 (0.03)
Math	D2	2015	4,529	<b>-0.10**</b> (0.04)	<b>-0.12***</b> (0.04)	-0.07 (0.05)	-0.07 (0.05)	0.04 (0.04)	0.04 (0.04)
		2019	5,669	-0.06 (0.05)	-0.05 (0.05)	-0.03 (0.04)	-0.03 (0.04)	0.02 (0.03)	0.03 (0.03)
		2015-19	10,198	<b>-0.07*</b> (0.04)	<b>-0.08**</b> (0.04)	-0.04 (0.03)	-0.05 (0.03)	0.02 (0.02)	0.03 (0.02)
	D3	2015	4,529	-0.08 (0.05)	-0.09 (0.06)	-0.08 (0.06)	-0.08 (0.06)	0.01 (0.04)	0.01 (0.03)
		2019	5,669	-0.05 (0.04)	-0.03 (0.05)	-0.03 (0.05)	-0.03 (0.05)	0.00 (0.03)	0.01 (0.03)
		2015-19	10,198	-0.05 (0.03)	-0.06 (0.04)	-0.04 (0.04)	-0.05 (0.04)	0.00 (0.02)	0.01 (0.02)
Science	D2	2015	4,529	<b>-0.11*</b> (0.06)	<b>-0.13**</b> (0.06)	<b>-0.09**</b> (0.04)	<b>-0.10**</b> (0.04)	0.02 (0.04)	0.02 (0.04)
		2019	5,669	-0.05 (0.04)	-0.04 (0.04)	-0.08* (0.04)	<b>-0.08*</b> (0.04)	-0.04 (0.03)	-0.03 (0.03)
		2015-19	10,198	<b>-0.07**</b> (0.03)	<b>-0.07**</b> (0.03)	<b>-0.08**</b> (0.03)	<b>-0.08***</b> (0.03)	-0.02 (0.03)	-0.01 (0.03)
	D3	2015	4,529	<b>-0.11*</b> (0.06)	<b>-0.12**</b> (0.06)	<b>-0.11**</b> (0.05)	<b>-0.12**</b> (0.05)	-0.01 (0.03)	0.00 (0.03)
		2019	5,669	-0.06 (0.04)	-0.04 (0.05)	<b>-0.09**</b> (0.05)	<b>-0.09*</b> (0.05)	-0.04 (0.03)	-0.04 (0.03)
		2015-19	10,198	<b>-0.07*</b> (0.04)	<b>-0.07*</b> (0.04)	<b>-0.09**</b> (0.04)	<b>-0.10***</b> (0.04)	-0.03 (0.03)	-0.03 (0.03)

Notes: see end of Table 3.

Table 3: Continued

Score	D	Year	N	ATET Medium vs Low		ATET High vs Low		ATET High vs Medium	
				T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
<b>Norway Grade 5</b>									
Reading	D2	2016	3,517	0.08 (0.05)	0.08 (0.05)	<b>0.15***</b> (0.05)	<b>0.16***</b> (0.06)	0.07 (0.04)	<b>0.07*</b> (0.04)
	D3	2016	3,517	0.00 (0.05)	0.01 (0.05)	<b>0.13**</b> (0.06)	<b>0.12*</b> (0.06)	<b>0.09**</b> (0.04)	<b>0.11***</b> (0.04)
Math	D2	2015	1,795	0.00 (0.16)	0.11 (0.17)	0.03 (0.13)	0.13 (0.13)	0.05 (0.07)	0.05 (0.07)
		2019	3,334	-0.02 (0.10)	0.00 (0.14)	0.15 (0.13)	<b>0.18*</b> (0.10)	<b>0.14**</b> (0.06)	<b>0.13**</b> (0.05)
		2015-19	5,129	0.00 (0.10)	0.04 (0.09)	0.11 (0.11)	0.14* (0.09)	0.12** (0.05)	0.10** (0.04)
	D3	2015	1,795	0.19 (0.21)	<b>0.34**</b> (0.17)	0.21 (0.19)	<b>0.38**</b> (0.18)	0.04 (0.06)	0.04 (0.06)
		2019	3,334	0.02 (0.18)	0.09 (0.12)	0.19 (0.23)	<b>0.30**</b> (0.12)	<b>0.13**</b> (0.06)	<b>0.12**</b> (0.05)
		2015-19	5,129	0.12 (0.14)	0.15 (0.10)	0.18 (0.19)	<b>0.30***</b> (0.11)	<b>0.11**</b> (0.05)	<b>0.08**</b> (0.04)
Science	D2	2015	1,795	0.28** (0.12)	0.37*** (0.13)	0.23** (0.11)	0.26** (0.11)	-0.03 (0.06)	-0.05 (0.07)
		2019	3,334	0.11 (0.10)	0.15 (0.15)	<b>0.26**</b> (0.12)	<b>0.31***</b> (0.11)	<b>0.15**</b> (0.06)	<b>0.13**</b> (0.06)
		2015-19	5,129	<b>0.19**</b> (0.08)	<b>0.22**</b> (0.10)	<b>0.26***</b> (0.10)	<b>0.28***</b> (0.09)	<b>0.11**</b> (0.05)	0.06 (0.04)
	D3	2015	1,795	<b>0.33*</b> (0.18)	<b>0.42**</b> (0.18)	0.32 (0.19)	<b>0.39*</b> (0.20)	0.00 (0.05)	-0.01 (0.06)
		2019	3,334	0.10 (0.15)	<b>0.24*</b> (0.13)	0.24 (0.21)	<b>0.47***</b> (0.11)	<b>0.16***</b> (0.06)	<b>0.13**</b> (0.05)
		2015-19	5,129	<b>0.22*</b> (0.13)	<b>0.28**</b> (0.11)	<b>0.27*</b> (0.16)	<b>0.43***</b> (0.10)	<b>0.12***</b> (0.04)	<b>0.08**</b> (0.04)
<b>Sweden</b>									
Reading	D2	2016	3,446	0.09 (0.09)	0.12 (0.09)	0.13 (0.08)	0.14* (0.08)	0.04 (0.05)	0.03 (0.05)
	D3	2016	3,446	<b>0.19*</b> (0.10)	<b>0.23**</b> (0.09)	<b>0.25**</b> (0.11)	<b>0.22**</b> (0.10)	0.04 (0.04)	0.03 (0.04)
Math	D2	2015	6,288	0.00 (0.08)	0.05 (0.09)	0.04 (0.08)	0.03 (0.08)	0.03 (0.06)	0.03 (0.05)
		2019	4,049	0.01 (0.09)	0.02 (0.11)	0.11 (0.07)	-0.04 (0.11)	0.08 (0.05)	0.07 (0.05)
		2015-19	10,337	0.01 (0.06)	0.05 (0.07)	0.08 (0.06)	-0.01 (0.07)	0.06 (0.04)	0.05 (0.04)
	D3	2015	6,288	0.02 (0.10)	0.06 (0.11)	0.06 (0.10)	0.02 (0.11)	0.03 (0.05)	0.03 (0.05)
		2019	4,049	0.09 (0.10)	-0.04 (0.13)	<b>0.21**</b> (0.09)	-0.06 (0.15)	0.07 (0.05)	0.06 (0.05)
		2015-19	10,337	0.06 (0.08)	0.02 (0.09)	0.16** (0.08)	-0.04 (0.10)	0.05 (0.04)	0.05 (0.04)
Science	D2	2015	6,288	-0.02 (0.08)	0.07 (0.09)	0.00 (0.08)	0.01 (0.11)	0.04 (0.06)	0.04 (0.06)
		2019	4,049	<b>0.24**</b> (0.11)	0.19 (0.13)	<b>0.30***</b> (0.09)	<b>0.23*</b> (0.12)	0.04 (0.05)	0.04 (0.05)
		2015-19	10,337	0.10 (0.07)	<b>0.16**</b> (0.08)	<b>0.16**</b> (0.06)	0.07 (0.10)	0.05 (0.04)	0.04 (0.04)
	D3	2015	6,288	0.11 (0.11)	<b>0.22*</b> (0.13)	0.10 (0.10)	0.09 (0.15)	0.01 (0.06)	0.01 (0.06)
		2019	4,049	<b>0.38**</b> (0.15)	0.11 (0.14)	<b>0.47***</b> (0.15)	0.23 (0.15)	0.06 (0.04)	0.07 (0.05)
		2015-19	10337	<b>0.22**</b> (0.10)	<b>0.18*</b> (0.10)	<b>0.29***</b> (0.09)	0.09 (0.13)	0.04 (0.04)	0.04 (0.04)

Notes: see end of Table 3.

Table 3: Continued

Score	D	Year	N	ATET Medium vs Low		ATET High vs Low		ATET High vs Medium	
				T <sub>ML M</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM M</sub>
<b>Belgium (Flemish)</b>									
Reading	D2	2016	4,204	0.14*** (0.05)	0.17*** (0.05)	0.20*** (0.05)	0.22*** (0.05)	0.08*** (0.03)	0.08*** (0.03)
	D3	2016	4,204	0.11 (0.07)	0.13* (0.07)	0.19** (0.09)	0.20*** (0.05)	0.10*** (0.03)	0.10*** (0.03)
Math	D2	2015	4,398	0.10* (0.05)	0.09 (0.06)	0.18*** (0.05)	0.19*** (0.06)	0.11*** (0.03)	0.10*** (0.03)
		2019	3,668	0.09* (0.05)	0.11** (0.05)	0.18*** (0.06)	0.19*** (0.05)	0.11*** (0.03)	0.08** (0.03)
		2015-19	8,066	0.10** (0.04)	0.09** (0.04)	0.18*** (0.04)	0.18*** (0.04)	0.11*** (0.02)	0.09*** (0.02)
	D3	2015	4,398	0.14** (0.06)	0.16** (0.07)	0.21*** (0.08)	0.25*** (0.06)	0.12*** (0.03)	0.10*** (0.03)
		2019	3,668	0.21*** (0.07)	0.15** (0.07)	0.34*** (0.09)	0.26*** (0.08)	0.11*** (0.03)	0.09*** (0.03)
		2015-19	8,066	0.17*** (0.04)	0.15*** (0.05)	0.28*** (0.06)	0.25*** (0.05)	0.11*** (0.02)	0.10*** (0.02)
Science	D2	2015	4,398	0.18*** (0.05)	0.19*** (0.05)	0.24*** (0.05)	0.26*** (0.05)	0.11*** (0.03)	0.09*** (0.03)
		2019	3,668	0.13** (0.06)	0.13** (0.06)	0.23*** (0.05)	0.22*** (0.05)	0.11** (0.04)	0.08** (0.04)
		2015-19	8,066	0.16*** (0.04)	0.16*** (0.04)	0.23*** (0.04)	0.24*** (0.03)	0.11*** (0.02)	0.09*** (0.02)
	D3	2015	4,398	0.24*** (0.07)	0.26*** (0.06)	0.31*** (0.08)	0.35*** (0.06)	0.12*** (0.03)	0.10*** (0.03)
		2019	3,668	0.31*** (0.11)	0.23*** (0.09)	0.46*** (0.12)	0.32*** (0.08)	0.11*** (0.03)	0.10** (0.04)
		2015-19	8,066	0.27*** (0.06)	0.25*** (0.05)	0.39*** (0.08)	0.33*** (0.05)	0.12*** (0.02)	0.10*** (0.02)
<b>France</b>									
Reading	D2	2016	3,757	0.14*** (0.04)	0.14*** (0.04)	0.29*** (0.05)	0.25*** (0.05)	0.13*** (0.04)	0.12*** (0.04)
	D3	2016	3,757	0.18*** (0.05)	0.17*** (0.05)	0.34*** (0.06)	0.29*** (0.06)	0.13*** (0.04)	0.12*** (0.04)
Math	D2	2015	3,581	0.12 (0.07)	0.13* (0.07)	0.21*** (0.08)	0.21*** (0.07)	0.09 (0.05)	0.08 (0.05)
		2019	4,680	0.08* (0.05)	0.11** (0.05)	0.16*** (0.05)	0.14** (0.06)	0.04 (0.03)	0.06* (0.03)
		2015-19	8,261	0.10** (0.04)	0.12*** (0.04)	0.17*** (0.04)	0.16*** (0.05)	0.05** (0.03)	0.06** (0.03)
	D3	2015	3,581	0.08 (0.06)	0.09 (0.06)	0.18** (0.08)	0.18** (0.08)	0.10** (0.05)	0.09* (0.05)
		2019	4,680	0.13** (0.06)	0.14** (0.07)	0.19*** (0.07)	0.18** (0.07)	0.05 (0.03)	0.06* (0.03)
		2015-19	8,261	0.12** (0.05)	0.12** (0.05)	0.19*** (0.05)	0.18*** (0.06)	0.06** (0.03)	0.07** (0.03)
Science	D2	2015	3,581	0.06 (0.07)	0.07 (0.07)	0.16* (0.08)	0.13 (0.08)	0.08* (0.05)	0.06 (0.05)
		2019	4,680	0.12** (0.05)	0.15*** (0.05)	0.20*** (0.06)	0.20*** (0.07)	0.05 (0.04)	0.07 (0.04)
		2015-19	8,261	0.10** (0.04)	0.13*** (0.04)	0.18*** (0.05)	0.17*** (0.06)	0.05 (0.03)	0.06* (0.03)
	D3	2015	3,581	0.09 (0.08)	0.11 (0.08)	0.16 (0.10)	0.17* (0.10)	0.09* (0.05)	0.07 (0.05)
		2019	4,680	0.21*** (0.06)	0.21*** (0.07)	0.28*** (0.08)	0.27*** (0.08)	0.06 (0.04)	0.07 (0.04)

Notes: see end of Table 3.



Table 3: End

Score	D	Year	N	ATET Medium vs Low		ATET High vs Low		ATET High vs Medium	
				T <sub>ML JM</sub>	T <sub>ML L</sub>	T <sub>HL H</sub>	T <sub>HL L</sub>	T <sub>HM H</sub>	T <sub>HM JM</sub>
<b>Ontario</b>									
Reading	D2	2016	3051	0.09 (0.07)	0.07 (0.07)	0.05 (0.05)	0.08 (0.05)	-0.03 (0.07)	-0.03 (0.07)
	D3	2016	3051	0.09 (0.07)	0.09 (0.06)	0.07 (0.07)	0.11 (0.07)	-0.01 (0.05)	0.01 (0.05)
Math	D2	2015	3366	0.03 (0.06)	0.00 (0.06)	0.06 (0.05)	0.01 (0.05)	0.03 (0.06)	0.03 (0.06)
		2019	2037	0.06 (0.06)	0.01 (0.07)	0.12 (0.09)	0.12 (0.10)	0.05 (0.09)	0.05 (0.08)
		2015/19	5403	0.04 (0.04)	0.01 (0.05)	<b>0.09*</b> (0.05)	0.06 (0.06)	0.04 (0.05)	0.04 (0.05)
	D3	2015	3366	-0.03 (0.04)	-0.04 (0.04)	0.05 (0.05)	-0.02 (0.06)	0.06 (0.05)	0.06 (0.05)
		2019	2037	0.03 (0.07)	0.02 (0.08)	0.14 (0.11)	0.12 (0.12)	0.08 (0.08)	0.09 (0.07)
		2015-19	5403	0.00 (0.04)	-0.02 (0.04)	0.08 (0.06)	0.04 (0.06)	0.07 (0.05)	0.07 (0.05)
Science	D2	2015	3366	0.05 (0.05)	0.04 (0.05)	0.03 (0.04)	0.00 (0.05)	-0.02 (0.05)	-0.02 (0.05)
		2019	2037	0.06 (0.07)	0.05 (0.07)	0.07 (0.07)	0.05 (0.07)	0.00 (0.10)	-0.01 (0.08)
		2015-19	5403	0.05 (0.04)	0.05 (0.04)	0.05 (0.04)	0.02 (0.04)	-0.01 (0.05)	-0.02 (0.05)
	D3	2015	3366	0.02 (0.04)	0.02 (0.04)	0.03 (0.05)	-0.02 (0.05)	0.00 (0.04)	0.00 (0.04)
		2019	2037	0.01 (0.06)	-0.02 (0.07)	0.08 (0.09)	0.02 (0.09)	0.04 (0.07)	0.04 (0.07)
		2015-19	5403	0.01 (0.04)	0.01 (0.04)	0.04 (0.05)	0.00 (0.05)	0.02 (0.04)	0.02 (0.04)
<b>Québec</b>									
Reading	D2	2016	2506	0.09 (0.07)	0.07 (0.07)	<b>0.11**</b> (0.04)	0.09** (0.04)	0.01 (0.06)	0.01 (0.07)
	D3	2016	2506	0.05 (0.06)	0.05 (0.06)	<b>0.10**</b> (0.05)	<b>0.10**</b> (0.04)	0.05 (0.05)	0.04 (0.05)
Math	D2	2015	2256	-0.04 (0.05)	-0.07 (0.06)	<b>0.10*</b> (0.05)	0.04 (0.05)	<b>0.13***</b> (0.04)	<b>0.14***</b> (0.04)
		2019	1764	0.04 (0.06)	0.06 (0.06)	<b>0.13***</b> (0.05)	<b>0.16***</b> (0.05)	0.08 (0.05)	<b>0.10*</b> (0.05)
		2015-19	4020	-0.01 (0.04)	-0.02 (0.04)	<b>0.11***</b> (0.04)	<b>0.08**</b> (0.04)	<b>0.11***</b> (0.03)	<b>0.11***</b> (0.03)
	D3	2015	2256	-0.02 (0.05)	-0.01 (0.06)	0.09 (0.06)	0.05 (0.06)	<b>0.11***</b> (0.04)	<b>0.09**</b> (0.04)
		2019	1764	0.01 (0.06)	-0.01 (0.06)	<b>0.13**</b> (0.06)	<b>0.15**</b> (0.07)	<b>0.10**</b> (0.04)	<b>0.13***</b> (0.04)
		2015/19	4020	0.00 (0.04)	-0.01 (0.04)	<b>0.11**</b> (0.04)	<b>0.09*</b> (0.05)	<b>0.11***</b> (0.03)	<b>0.10***</b> (0.03)
Science	D2	2015	2256	-0.01 (0.06)	-0.06 (0.07)	0.01 (0.06)	-0.05 (0.06)	0.01 (0.06)	0.02 (0.06)
		2019	1764	0.03 (0.06)	0.04 (0.05)	0.05 (0.05)	0.07 (0.05)	0.01 (0.05)	0.03 (0.06)
		2015-19	4020	0.00 (0.04)	-0.02 (0.05)	0.03 (0.04)	0.00 (0.04)	0.00 (0.04)	0.01 (0.04)
	D3	2015	2256	0.01 (0.06)	0.00 (0.07)	0.00 (0.07)	-0.04 (0.07)	0.00 (0.05)	-0.02 (0.05)
		2019	1764	0.01 (0.06)	0.00 (0.06)	0.06 (0.06)	0.07 (0.06)	0.03 (0.05)	0.05 (0.06)
		2015-19	4020	0.01 (0.04)	0.00 (0.05)	0.03 (0.05)	0.00 (0.05)	0.01 (0.04)	0.00 (0.04)

Notes: Coefficient (Std. Err.) Indicate estimated parameters. Preschool durations used are: D2: (1) ≤ 2 years; (2) 3 years; (3) ≥ 4 years; D3: (1) ≤ 1 year; (2) 1 to ≤ 3 years; (3) ≥ 4 years. Statistical significance: If a p-value is less than 0.05, one star (\*); If less than 0.01, two stars (\*\*); If less than 0.001, three stars (\*\*\*)

**Table 4: Average treatments effects (ATE) estimates (std. err.) of the preschool programs participation types on z-scores, reading (PIRLS 2016) math, science (TIMSS 2015, 2019), and years pooling by jurisdiction**

Score	Year	N	T21 Early vs No	T31 Late vs No	T41 Both vs No	T32 Late vs Early	T42 Both vs Early	T43 Both vs Late
Finland								
Reading	2016	3,690	-0.06 (0.38)	-0.03 (0.36)	-0.04 (0.41)	0.07 (0.10)	-0.01 (0.10)	-0.04 (0.03)
Math	2015	4,536	0.04 (0.21)	0.10 (0.20)	0.13 (0.21)	<b>0.12**</b> (0.06)	<b>0.13**</b> (0.05)	-0.01 (0.04)
	2019	4,798	-0.21 (0.47)	-0.07 (0.30)	-0.13 (0.37)	0.07 (0.20)	0.10 (0.17)	-0.01 (0.04)
	2015-19	9,334	-0.06 (0.16)	0.02 (0.18)	-0.04 (0.25)	0.11 (0.11)	0.11 (0.11)	-0.01 (0.03)
Science	2015	4,536	-0.15 (0.27)	-0.05 (0.26)	-0.05 (0.30)	<b>0.14**</b> (0.06)	<b>0.12**</b> (0.05)	-0.02 (0.04)
	2019	4,798	-0.28 (0.31)	-0.08 (0.19)	-0.21 (0.21)	<b>0.25*</b> (0.13)	<b>0.22*</b> (0.11)	-0.07 (0.04)
	2015-19	9,334	-0.17 (0.23)	0.05 (0.20)	-0.07 (0.20)	<b>0.21**</b> (0.08)	<b>0.18**</b> (0.08)	<b>-0.05*</b> (0.03)
Norway Grade 5								
Reading	2016	3,465	0.12 (0.21)	-0.01 (0.16)	0.13 (0.19)	-0.15 (0.12)	0.05 (0.10)	0.09 (0.08)
Math	2015	1,794	-0.22 (0.74)	-0.01 (0.55)	-0.10 (0.71)	0.04 (0.09)	<b>0.22***</b> (0.05)	<b>0.18**</b> (0.08)
	2019	2,441	-0.04 (0.25)	0.17 (0.19)	0.17 (0.25)	0.05 (0.14)	<b>0.24***</b> (0.08)	<b>0.18**</b> (0.08)
	2015-19	4,235	-0.09 (0.18)	0.08 (0.21)	0.14 (0.18)	0.08 (0.08)	<b>0.24***</b> (0.06)	<b>0.20***</b> (0.07)
Science	2015	1,794	-0.25 (0.55)	-0.06 (0.46)	-0.18 (0.50)	0.00 (0.08)	<b>0.12**</b> (0.05)	<b>0.12**</b> (0.06)
	2019	2,441	-0.06 (0.24)	0.21 (0.23)	0.19 (0.27)	0.07 (0.16)	<b>0.26***</b> (0.07)	<b>0.28***</b> (0.09)
	2015-19	4,235	-0.05 (0.17)	0.05 (0.24)	0.14 (0.17)	0.05 (0.09)	<b>0.22***</b> (0.06)	<b>0.25***</b> (0.07)
Sweden								
Reading	2016	3,130	-0.01 (0.39)	0.03 (0.27)	-0.16 (0.46)	0.15 (0.17)	0.01 (0.13)	0.05 (0.07)
Math	2015	6,292	-0.20 (0.12)	-0.06 (0.12)	-0.01 (0.13)	0.07 (0.07)	<b>0.16***</b> (0.05)	0.05 (0.07)
	2019	3,475	<b>0.46**</b> (0.22)	-0.02 (0.14)	<b>0.27*</b> (0.14)	<b>-0.26*</b> (0.14)	-0.10 (0.09)	<b>0.20***</b> (0.07)
Science	2015	6,292	-0.08 (0.13)	0.14 (0.17)	0.09 (0.16)	0.10 (0.10)	<b>0.15***</b> (0.05)	-0.02 (0.07)
	2019	3,475	<b>0.67***</b> (0.23)	0.33 (0.27)	<b>0.55*</b> (0.28)	-0.22 (0.19)	-0.01 (0.09)	<b>0.21***</b> (0.08)
Belgium Flemish								
Reading	2016	3,756	-0.09 (0.20)	0.11 (0.12)	0.16 (0.13)	0.21 (0.24)	0.08 (0.22)	<b>0.08***</b> (0.03)
Math	2015	4,420	-0.07 (0.12)	0.08 (0.11)	0.09 (0.16)	<b>0.14**</b> (0.07)	<b>0.27***</b> (0.09)	<b>0.06*</b> (0.03)
	2019	3,423	0.06 (0.34)	0.20 (0.13)	<b>0.40**</b> (0.14)	0.05 (0.20)	0.28 (0.41)	<b>0.08**</b> (0.04)
	2015-19	7,843	-0.05 (0.11)	0.14 (0.08)	<b>0.27**</b> (0.11)	0.11 (0.07)	<b>0.25*</b> (0.14)	<b>0.07***</b> (0.02)

Notes: see end of Table 4.

Table 4: End

Score	Year	N	T21 Early vs no	T31 Late vs no	T41 Both vs No	T32 Late vs Early	T42 Both vs Early	T43 Both vs Late
Belgium Flemish end								
Science	2015	4,420	0.01 (0.15)	0.15 (0.12)	0.19 (0.17)	<b>0.16*</b> (0.08)	<b>0.26***</b> (0.09)	<b>0.06*</b> (0.03)
	2019	3,423	-0.04 (0.51)	0.28 (0.20)	<b>0.51***</b> (0.15)	0.16 (0.28)	<b>0.41*</b> (0.24)	<b>0.08*</b> (0.05)
	2015/19	7,843	0.01 (0.14)	<b>0.20**</b> (0.09)	<b>0.38***</b> (0.11)	<b>0.20**</b> (0.10)	<b>0.38***</b> (0.14)	<b>0.07**</b> (0.03)
France								
Reading	2016	3,404	0.10 (0.12)	<b>0.16**</b> (0.08)	<b>0.25***</b> (0.08)	0.06 (0.11)	0.14 (0.11)	<b>0.08*</b> (0.04)
Math	2015	3,616	-0.01 (0.13)	0.06 (0.11)	0.10 (0.12)	0.08 (0.06)	<b>0.16*</b> (0.09)	0.07 (0.06)
	2019	4,207	-0.05 (0.19)	0.11 (0.10)	0.11 (0.11)	0.17 (0.13)	0.18 (0.14)	-0.01 (0.04)
	2015/19	7,823	-0.03 (0.10)	0.08 (0.07)	0.10 (0.07)	0.14 (0.09)	<b>0.18*</b> (0.11)	0.02 (0.03)
Science	2015	3,616	0.01 (0.14)	0.04 (0.13)	0.10 (0.14)	0.03 (0.05)	<b>0.12*</b> (0.07)	0.08 (0.05)
	2019	4,207	-0.15 (0.25)	0.13 (0.16)	0.12 (0.16)	0.29 (0.17)	0.28 (0.18)	0.00 (0.04)
	2015/19	7,823	-0.03 (0.12)	0.10 (0.10)	0.11 (0.11)	<b>0.20*</b> (0.11)	<b>0.23*</b> (0.12)	0.02 (0.03)
Ontario								
Reading	2016	2,727	-0.08 (0.11)	<b>0.12*</b> (0.06)	<b>0.14*</b> (0.07)	<b>0.19*</b> (0.10)	0.17 (0.11)	-0.01 (0.05)
Math	2015	3,166	0.06 (0.06)	-0.05 (0.05)	0.01 (0.05)	<b>-0.12**</b> (0.05)	-0.05 (0.07)	0.04 (0.04)
	2019	1,924	0.04 (0.09)	<b>0.15**</b> (0.07)	<b>0.25***</b> (0.08)	0.11 (0.09)	<b>0.20*</b> (0.10)	<b>0.10**</b> (0.04)
Science	2015	3,166	0.04 (0.06)	0.00 (0.04)	0.01 (0.06)	-0.03 (0.05)	-0.03 (0.05)	-0.01 (0.05)
	2019	1,924	0.12 (0.13)	0.13 (0.09)	<b>0.18*</b> (0.09)	0.03 (0.10)	0.08 (0.10)	0.04 (0.05)
	2015/19	5,090	0.05 (0.06)	0.06 (0.05)	0.10 (0.06)	-0.01 (0.06)	0.03 (0.06)	0.02 (0.04)
Québec								
Reading	2016	2,250	<b>0.20*</b> (0.10)	<b>0.20***</b> (0.05)	<b>0.22***</b> (0.06)	0.06 (0.16)	0.05 (0.14)	0.01 (0.04)
Math	2015	2,194	-0.10 (0.12)	-0.02 (0.07)	-0.04 (0.08)	0.12 (0.10)	0.10 (0.09)	-0.03 (0.07)
	2019	1,691	0.08 (0.14)	0.15 (0.11)	<b>0.22*</b> (0.12)	0.11 (0.07)	<b>0.15**</b> (0.07)	0.05 (0.06)
	2015/19	3,885	-0.02 (0.08)	0.05 (0.06)	0.07 (0.07)	0.09 (0.06)	<b>0.10*</b> (0.06)	0.01 (0.04)
Science	2015	2,194	<b>-0.17*</b> (0.10)	-0.06 (0.07)	-0.09 (0.08)	<b>0.16*</b> (0.08)	0.11 (0.09)	-0.05 (0.06)
	2019	1,691	0.09 (0.15)	0.19 (0.12)	0.19 (0.12)	0.13 (0.09)	0.11 (0.08)	-0.01 (0.05)
	2015/19	3,885	-0.04 (0.09)	0.05 (0.06)	0.04 (0.07)	<b>0.13*</b> (0.06)	0.10 (0.06)	-0.03 (0.04)

Notes: The four types of preschool participation are: 1. No: the child did not participate in any program; 2. Early: the child participate in an early program (for less than 3 years-old); 3. Late: the child participate in a later program (for 3 years-old or more); 4. Both: the Child participates in both programs.

**Table 5: Average multivalued treatments effects of treated, ATET estimates (std. err.) of three parental education's levels on reading z-scores, by jurisdiction and years pooling, WLS-DR-ordered logit, PIRLS 2011+2016**

Jurisdiction Year	Education levels (N=observations)	ATET Medium vs Low		ATET High vs Low		ATET High vs Medium	
		T <sub>ML JM</sub>	T <sub>ML JL</sub>	T <sub>HL JH</sub>	T <sub>HL JL</sub>	T <sub>HM JH</sub>	T <sub>HM JM</sub>
Denmark	Educ1 (N=1,001)	0.49* (0.26)	0.38 (0.29)	0.65*** (0.19)	0.66*** (0.21)	0.30*** (0.12)	0.28*** (0.09)
2011&2016	Educ2 (N=1,617)	0.37 (0.35)	0.41 (0.03)	0.56* (0.30)	0.38 (0.27)	-0.08 (0.18)	0.11 (0.08)
N=6,676	Educ3 (N=4,058)	0.19 (0.15)	0.05 (0.31)	0.46** (0.19)	0.36** (0.31)	0.39*** (0.13)	0.13* (0.07)
Finland	Educ1 (N=2,283)	-0.11 (0.08)	-0.09 (0.08)	-0.11* (0.06)	-0.09 (0.06)	0.03 (0.08)	0.02 (0.08)
2011&2016	Educ2 (N=2,083)	-0.10 (0.08)	-0.09 (0.08)	-0.11* (0.08)	0.07 (0.07)	0.02 (0.07)	0.01 (0.06)
N=8,841	Educ3 (N=4,475)	-0.18** (0.07)	-0.15** (0.02)	-0.12** (0.08)	-0.14 (0.06)	0.01 (0.05)	0.02 (0.05)
Norway	Educ1 (N=909)	0.10 (0.15)	-0.09 (0.19)	0.08 (0.12)	-0.07 (0.12)	-0.02 (0.13)	-0.06 (0.11)
2011&2016	Educ2 (N=1,393)	-0.20 (0.18)	0.31 (0.20)	-0.06 (0.10)	0.02 (0.09)	0.18 (0.15)	-0.03 (0.09)
N=6,376	Educ3 (N=4,074)	-0.05 (0.14)	0.20 (0.13)	0.09 (0.08)	-0.08 (0.07)	0.22 (0.08)	-0.14 (0.07)
Sweden	Educ1 (N=2,412)	0.15 (0.13)	0.20 (0.15)	0.08 (0.12)	0.17 (0.11)	-0.17 (0.10)	0.01 (0.07)
2011&2016	Educ2 (N=1,868)	-0.03 (0.18)	-0.05 (0.14)	0.04 (0.09)	-0.10 (0.11)	0.10 (0.11)	-0.04 (0.08)
N=7,593	Educ3 (N=3,317)	-0.25 (0.36)	0.12 (0.22)	0.20 (0.21)	0.29 (0.19)	0.09 (0.11)	0.21*** (0.08)
France	Educ1 (N=3,545)	-0.21 (0.19)	0.09 (0.11)	-0.03 (0.13)	0.13* (0.08)	0.04 (0.05)	0.02 (0.05)
2011&2016	Educ2 (N=1,475)	0.17 (0.17)	0.3 (0.19)	0.12 (0.14)	0.26* (0.13)	-0.04 (0.08)	0.03 (0.07)
N=7,652	Educ3 (N=2,632)	0.33 (0.84)	0.2 (0.18)	0.07 (0.4)	0.38*** (0.09)	0.12 (0.10)	0.10 (0.08)
Belgium Fl.	Educ1 (N=2,412)	0.11 (0.35)	0.07 (0.21)	0.27 (0.22)	0.04 (0.15)	0.06 (0.08)	0.11 (0.07)
2006&2016	Educ2 (N=1,878)	0.16* (0.18)	-0.24** (0.19)	0.24* (0.14)	0.07 (0.12)	0.14 (0.10)	0.10 (0.09)
N=7,667	Educ3 (N=3,377)	0.15 (0.17)	0.03 (0.17)	0.19 (0.15)	0.18 (0.13)	0.11 (0.07)	0.06 (0.06)
Belgium Fr.	Educ1 (N=1,950)	-0.28* (0.15)	-0.14 (0.17)	-0.16 (0.14)	-0.17 (0.13)	-0.01 (0.08)	0.03 (0.07)
2011&2016	Educ2 (N=803)	0.50 (0.38)	0.05 (0.28)	0.81** (0.38)	0.09 (0.20)	0.29* (0.16)	0.16 (0.13)
N=6,375	Educ3 (N=3,622)	0.04 (0.27)	0.53** (0.24)	0.06 (0.17)	0.64*** (0.23)	0.09 (0.07)	0.21*** (0.07)
Ontario	Educ1 (N=712)	0.29** (0.13)	0.22* (0.12)	0.38** (0.18)	0.37** (0.17)	0.07 (0.18)	0.05 (0.13)
2011&2016	Educ2 (N=2,165)	0.02 (0.09)	0.06 (0.08)	0.02 (0.09)	0.03 (0.09)	-0.03 (0.09)	0.03 (0.09)
N=6,421	Educ3 (N=3,544)	0.09 (0.07)	0.07 (0.08)	0.15* (0.08)	0.17** (0.09)	0.11* (0.06)	0.10* (0.05)
Québec	Educ1 (N=689)	0.16* (0.09)	0.17** (0.08)	0.11 (0.13)	0.23* (0.14)	-0.02 (0.18)	0.08 (0.16)
2011&2016	Educ2 (N=2,247)	0.00 (0.0)	-0.03 (0.07)	0.07 (0.07)	0.12 (0.08)	0.17 (0.10)	0.14 (0.09)
N=5,977	Educ3 (N=3,041)	0.01 (0.05)	0.03 (0.05)	0.15*** (0.06)	0.21*** (0.06)	0.07 (0.07)	0.18** (0.08)

Notes: Education levels (see Table C): Educ1 (Lower): ≤Secondary; Educ2 (Medium): Postsecondary; Educ3 (High): ≥University. Belgium Fl.: Belgium Flemish jurisdiction; Belgium Fr.: Belgium French jurisdiction.

## Supplementary statistical Tables

**Table A1: Number of students (N) and percentage distribution of parental education levels by preschool duration attendance levels in years, PIRLS 2006, 2011, 2016, and TIMSS 2015 and 2019**

	PIRLS			TIMSS		PIRLS			TIMSS	
	2006	2011	2016	2015	2019	2006	2011	2016	2015	2019
	Denmark					Finland				
Low preschool (<=1 year) N	216	191	51	235	37	NA	1,498	797	1,498	601
≤Secondary %	36	32	39	12	24	NA	36	36	36	35
Postsecondary %	24	19	12	29	37	NA	27	22	27	14
≥University %	40	49	49	59	40	NA	37	42	37	51
Middle preschool (1-2 years) N	440	501	434	1,243	48	NA	733	1,599	733	402
≤Secondary %	25	21	22	15	15	NA	28	28	28	27
Postsecondary %	33	32	27	30	26	NA	30	19	30	15
≥University %	42	46	51	55	59	NA	43	53	43	58
High preschool (=>3 years) N	2,661	3,201	2,539	1,740	2,509	NA	2,232	2,110	2,232	3,048
≤Secondary %	19	17	12	8	9	NA	26	20	26	24
Postsecondary %	30	24	22	24	17	NA	27	20	27	14
≥University %	52	59	66	69	73	NA	47	60	47	63
	Norway Grade 5 (# if Grade 4)					Sweden				
Low preschool (<=1 year) N	758#	347#	454	130	95	909	548	169	626	220
≤Secondary %	35	36	23	24	34	42	46	36	36	48
Postsecondary %	35	32	29	42	21	37	27	29	34	29
≥University %	31	32	48	35	45	22	27	35	31	22
Middle preschool (1-2 years) N	496#	412#	766	387	88	393	358	592	924	641
≤Secondary %	26	22	25	22	39	30	27	18	26	22
Postsecondary %	35	30	29	30	34	40	38	25	25	25
≥University %	39	48	46	48	28	30	35	57	50	53
High preschool (=>3 years) N	2,142#	2,017#	2,480	1,295	3,290	2,166	2,740	2,821	5,072	3,425
≤Secondary %	17	13	10	9	12	23	27	18	21	18
Postsecondary %	26	22	21	24	20	36	26	25	25	23
≥University %	58	65	69	67	68	41	47	57	54	59
	Belgium Flemish					Belgium French				
Low preschool (<=1 year) N	99	NA	193	629	149	220	120	66	NA	NA
≤Secondary %	55	NA	42	45	48	47	56	51	NA	NA
Postsecondary %	15	NA	29	22	15	41	11	17	NA	NA
≥University %	31	NA	29	33	36	12	32	32	NA	NA
Middle preschool (1-2 years) N	461	NA	1,54	1,337	207	664	583	187	NA	NA
≤Secondary %	46	NA	33	32	25	51	40	44	NA	NA
Postsecondary %	23	NA	25	25	25	40	16	11	NA	NA
≥University %	31	NA	42	44	49	9	44	46	NA	NA
High preschool (=>3 years) N	3,518	NA	2,782	2,861	3,570	2,937	2,383	3,413	NA	NA
≤Secondary %	40	NA	14	16	20	40	34	29	NA	NA
Postsecondary %	28	NA	17	17	17	51	13	13	NA	NA
≥University %	32	NA	69	67	63	10	54	58	NA	NA

Notes and Sources: See end of Table A1.

**Table A1: End**

	PIRLS			TIMSS						
	2006	2011	2016	2015	2019					
France										
Low preschool (<=1 year) N	160	151	532	562	348					
≤Secondary %	65	61	57	56	54					
Postsecondary %	11	13	19	19	18					
≥University %	24	26	25	25	28					
Middle preschool (1-2 years) N	739	807	1,579	2,372	257					
≤Secondary %	63	58	42	45	50					
Postsecondary %	15	19	24	23	15					
≥University %	21	23	34	33	35					
High preschool (=>3 years) N	2,864	2,969	950	1,127	4,624					
≤Secondary %	57	50	38	39	39					
Postsecondary %	16	18	20	23	23					
≥University %	27	32	42	38	38					
Ontario										
Low preschool (<=1 year) N	2,763	2,385	859	2,096	683					
≤Secondary %	24	19	19	18	12					
Postsecondary %	38	40	43	34	38					
≥University %	38	41	38	48	50					
Middle preschool (1-2 years) N	262	441	1,791	801	660					
≤Secondary %	16	12	10	8	10					
Postsecondary %	31	32	36	32	30					
≥University %	53	57	54	60	60					
High preschool (=>3 years) N	321	652	637	628	893					
≤Secondary %	10	9	6	5	5					
Postsecondary %	31	27	35	32	27					
≥University %	59	64	59	63	69					
Québec										
Low preschool (<=1 year) N	2,463	2,752	481	912	430					
≤Secondary %	21	15	15	19	19					
Postsecondary %	40	42	43	35	38					
≥University %	39	44	42	47	44					
Middle preschool (1-2 years) N	233	304	1,040	558	309					
≤Secondary %	13	13	10	10	12					
Postsecondary %	30	33	41	30	29					
≥University %	57	53	49	60	59					
High preschool (=>3 years) N	349	547	1,166	885	1,392					
≤Secondary %	16	14	6	6	7					
Postsecondary %	32	33	37	33	35					
≥University %	52	53	57	61	58					

Notes: If student's information on parental education and preschool duration are not missing. NA: jurisdiction did not participate in the year-survey. Attendance levels: Low: 0 to <1 years; Middle: 1-2 years; High: >3 years. Parental education levels: Lower: ≤Secondary; Middle: postsecondary; High: ≥University.

Source: Authors' computation from weighted PIRLS and TIMSS data sets.

**Table A2: Percentile distribution of students' reading z-scores, by preschool duration (in years), parental education levels, selected jurisdiction, PIRLS 2006-2011-2016**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25
Preschool duration levels in years										Parental education levels									
<b>Denmark 2006</b>	3,317	1.00	0.23	1.01	-0.36	0.34	0.92	1.43	1.27	<b>Denmark 2006</b>	3,317	1.00	0.23	1.01	-0.36	0.34	0.92	1.43	1.27
<=1 year	216	0.07	0.00	1.04	-0.67	0.10	0.73	1.31	1.40	≤Secondary	662	0.20	-0.18	1.06	-0.79	-0.04	0.53	1.04	1.33
1-2 years	440	0.13	0.16	1.01	-0.40	0.32	0.80	1.30	1.20	Postsecondary	946	0.29	0.17	1.00	-0.44	0.28	0.85	1.38	1.28
=>3 years	2,661	0.80	0.26	1.01	-0.32	0.37	0.95	1.46	1.27	≥University	1,709	0.52	0.43	0.95	-0.13	0.52	1.08	1.56	1.20
<b>Denmark 2011</b>	3,904	1.00	0.24	0.94	-0.34	0.31	0.89	1.38	1.24	<b>Denmark 2011</b>	3,904	1.00	0.24	0.94	-0.34	0.31	0.89	1.38	1.24
<=1 year	191	0.05	-0.14	0.99	-0.73	-0.07	0.51	1.14	1.24	≤Secondary	695	0.18	-0.14	0.97	-0.78	-0.07	0.51	1.04	1.29
1-2 years	501	0.13	0.12	0.93	-0.45	0.20	0.76	1.23	1.21	Postsecondary	962	0.25	0.12	0.93	-0.45	0.19	0.77	1.24	1.23
=>3 years	3,212	0.82	0.28	0.93	-0.30	0.35	0.93	1.42	1.23	≥University	2,247	0.58	0.42	0.89	-0.14	0.48	1.04	1.51	1.18
<b>Denmark 2016</b>	3,024	1.00	0.20	0.95	-0.36	0.27	0.85	1.33	1.21	<b>Denmark 2016</b>	3,024	1.00	0.20	0.95	-0.36	0.27	0.85	1.33	1.21
<=1 year	51	0.02	-0.55	1.28	-1.37	-0.48	0.34	1.00	1.71	≤Secondary	416	0.14	-0.29	0.94	-0.88	-0.19	0.39	0.81	1.27
1-2 years	434	0.14	-0.05	0.98	-0.66	0.07	0.63	1.11	1.29	Postsecondary	6,94	0.23	0.01	0.94	-0.58	0.08	0.65	1.13	1.23
=>3 years	2,539	0.84	0.26	0.92	-0.30	0.32	0.88	1.37	1.18	≥University	1,914	0.63	0.37	0.90	-0.16	0.43	0.98	1.45	1.15
<b>Finland 2011</b>	4,463	1.00	0.40	0.93	-0.18	0.44	1.04	1.56	1.22	<b>Finland 2011</b>	4,463	1.00	0.40	0.93	-0.18	0.44	1.04	1.56	1.22
<=1 year	1,498	0.34	0.41	0.95	-0.18	0.46	1.07	1.57	1.25	≤Secondary	1,233	0.28	0.07	0.92	-0.50	0.10	0.73	1.18	1.23
1-2 years	733	0.16	0.36	0.92	-0.23	0.39	1.00	1.49	1.23	Postsecondary	1,211	0.27	0.40	0.87	-0.16	0.41	0.99	1.52	1.15
=>3 years	2,232	0.50	0.41	0.93	-0.16	0.45	1.03	1.56	1.19	≥University	2,019	0.45	0.64	0.91	0.07	0.68	1.27	1.78	1.19
<b>Finland 2016</b>	4,506	1.00	0.45	0.92	-0.12	0.50	1.07	1.57	1.19	<b>Finland 2016</b>	4,506	1.00	0.45	0.92	-0.12	0.50	1.07	1.57	1.19
<=1 year	797	0.18	0.47	0.94	-0.11	0.51	1.12	1.60	1.23	≤Secondary	1,111	0.25	0.09	0.94	-0.46	0.14	0.72	1.20	1.19
1-2 years	1,599	0.35	0.40	0.92	-0.14	0.49	1.03	1.50	1.17	Postsecondary	903	0.20	0.30	0.91	-0.23	0.33	0.92	1.42	1.15
=>3 years	2,110	0.47	0.47	0.91	-0.1	0.52	1.09	1.59	1.19	≥University	2,492	0.55	0.67	0.85	0.15	0.71	1.25	1.72	1.10
<b>Norway 2006</b>	3,396	1.00	-0.50	0.96	-1.10	-0.42	0.17	0.67	1.27	<b>Norway 2006</b>	3,396	1.00	-0.50	0.96	-1.10	-0.42	0.17	0.67	1.27
<=1 year	758	0.22	-0.68	1.00	-1.33	-0.56	-0.02	0.53	1.31	≤Secondary	751	0.22	-0.90	0.99	-1.56	-0.79	-0.20	0.24	1.36
1-2 years	496	0.15	-0.64	0.98	-1.32	-0.54	0.09	0.59	1.41	Postsecondary	996	0.29	-0.61	0.93	-1.18	-0.52	0.00	0.47	1.18
=>3 years	2,142	0.63	-0.40	0.93	-0.96	-0.34	0.24	0.73	1.20	≥University	1,649	0.49	-0.25	0.90	-0.83	-0.19	0.40	0.84	1.23
<b>Norway 2011</b>	2,776	1.00	-0.47	0.90	-1.06	-0.41	0.15	0.63	1.21	<b>Norway 2011</b>	2,776	1.00	-0.47	0.90	-1.06	-0.41	0.15	0.63	1.21
<=1 year	347	0.13	-0.65	0.91	-1.28	-0.63	0.03	0.52	1.31	≤Secondary	468	0.17	-0.87	0.88	-1.43	-0.78	-0.26	0.22	1.17
1-2 years	412	0.15	-0.57	0.90	-1.15	-0.55	0.02	0.61	1.17	Postsecondary	619	0.22	-0.60	0.89	-1.14	-0.57	0.01	0.51	1.16
=>3 years	2,017	0.73	-0.42	0.90	-0.99	-0.35	0.19	0.65	1.19	≥University	1,689	0.61	-0.30	0.87	-0.87	-0.25	0.30	0.74	1.17
<b>Norway 2016</b>	3,812	1.00	-0.24	0.98	-0.84	-0.16	0.44	0.97	1.28	<b>Norway 2016</b>	3,812	1.00	-0.24	0.98	-0.84	-0.16	0.44	0.97	1.28
<=1 year	471	0.12	-0.41	0.97	-1.06	-0.32	0.27	0.84	1.32	≤Secondary	527	0.14	-0.69	0.97	-1.33	-0.63	-0.03	0.51	1.31
1-2 years	698	0.18	-0.43	0.99	-1.03	-0.31	0.25	0.78	1.28	Postsecondary	825	0.22	-0.45	0.94	-1.01	-0.35	0.21	0.68	1.22
=>3 years	2,643	0.69	-0.15	0.97	-0.75	-0.08	0.52	1.04	1.27	≥University	3,460	0.91	-0.05	0.95	-0.62	0.02	0.61	1.09	1.23
<b>Sweden 2006</b>	3,468	1.00	0.28	0.92	-0.27	0.33	0.90	1.40	1.17	<b>Sweden 2006</b>	3,468	1.00	0.28	0.92	-0.27	0.33	0.90	1.40	1.17
<=1 year	909	0.26	0.05	0.95	-0.50	0.16	0.70	1.17	1.20	≤Secondary	969	0.28	-0.06	0.94	-0.63	0.05	0.59	1.04	1.23
1-2 years	393	0.11	0.26	0.86	-0.23	0.29	0.83	1.29	1.06	Postsecondary	1,238	0.36	0.17	0.25	0.86	-0.29	0.31	0.84	-0.55
=>3 years	2,166	0.62	0.38	0.90	-0.17	0.41	0.99	1.49	1.16	≥University	1,261	0.36	0.59	0.87	0.04	0.61	1.18	1.67	1.14
<b>Sweden 2011</b>	3,646	1.00	0.09	0.94	-0.48	0.136	0.71	1.24	1.19	<b>Sweden 2011</b>	3,646	1.00	0.09	0.94	-0.48	0.14	0.71	1.24	1.19
<=1 year	548	0.15	-0.17	1.00	-0.79	-0.08	0.47	1.02	1.25	≤Secondary	1,090	0.30	-0.28	0.94	-0.88	-0.22	0.37	0.85	1.25
1-2 years	358	0.10	-0.01	0.83	-0.48	0.02	0.55	0.97	1.02	Postsecondary	996	0.27	0.02	0.86	-0.50	0.06	0.59	1.05	1.09
=>3 years	2,740	0.75	0.16	0.93	-0.41	0.19	0.79	1.31	1.20	≥University	1,560	0.43	0.39	0.89	-0.16	0.40	0.98	1.48	1.13

Notes and Sources: See end of Table A2.

Table A2: Continued

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25
Preschool duration in years										Parental education levels									
<b>Sweden 2016</b>	3,582	1.00	0.34	0.92	-0.22	0.41	0.97	1.45	1.19	<b>Sweden 2016</b>	3,582	1.00	0.34	0.92	-0.22	0.41	0.97	1.45	1.19
<=1 year	169	0.05	-0.09	0.98	-0.64	-0.07	0.45	1.16	1.09	≤Secondary	725	0.20	-0.11	0.93	-0.64	-0.02	0.52	0.98	1.16
1-2 years	592	0.17	0.27	0.92	-0.28	0.32	0.89	1.37	1.17	Postsecondary	918	0.26	0.17	0.91	-0.40	0.22	0.81	1.27	1.21
=>3 years	2,821	0.79	0.38	0.92	-0.18	0.46	1.01	1.47	1.18	≥University	1,939	0.54	0.58	0.85	0.08	0.64	1.16	1.60	1.08
<b>Belgium Fl. 2011</b>	4,078	1.00	0.22	0.80	-0.29	0.25	0.76	1.22	1.05	<b>Belgium Fl. 2011</b>	4,078	1.00	0.22	0.80	-0.29	0.25	0.76	1.22	1.05
<=1 year	99	0.02	-0.03	0.99	-0.58	0.11	0.67	1.08	1.25	≤Secondary	1,675	0.41	-0.11	0.77	-0.59	-0.08	0.41	0.86	1.00
1-2 years	461	0.11	0.12	0.86	-0.47	0.18	0.71	1.17	1.18	Postsecondary	1,097	0.27	0.35	0.71	-0.11	0.35	0.81	1.24	0.93
=>3 years	3,518	0.86	0.24	0.78	-0.26	0.26	0.77	1.23	1.03	≥University	1,306	0.32	0.54	0.75	0.08	0.58	1.02	1.48	0.94
<b>Belgium Fl. 2016</b>	4,465	1.00	-0.11	0.84	-0.66	-0.08	0.47	0.93	1.13	<b>Belgium Fl. 2016</b>	4,465	1.00	-0.11	0.84	-0.66	-0.08	0.47	0.93	1.13
<=1 year	193	0.04	-0.52	0.86	-1.09	-0.56	0.11	0.63	1.20	≤Secondary	981	0.22	-0.44	0.83	-0.98	-0.41	0.16	0.60	1.14
1-2 years	1,540	0.34	-0.27	0.85	-0.82	-0.22	0.32	0.79	1.14	Postsecondary	876	0.20	-0.35	0.78	-0.84	-0.31	0.18	0.62	1.02
=>3 years	2,732	0.61	0.01	0.80	-0.53	0.02	0.57	1.02	1.10	≥University	2,608	0.58	0.10	0.80	-0.42	0.13	0.66	1.08	1.08
<b>Belgium Fr. 2006</b>	3,821	1.00	-0.46	1.00	-1.11	-0.41	0.22	0.78	1.33	<b>Belgium Fr. 2006</b>	3,821	1.00	-0.46	1.00	-1.11	-0.41	0.22	0.78	1.33
<=1 year	220	0.06	-0.75	1.17	-1.56	-0.68	0.11	0.73	1.67	≤Secondary	1,566	0.41	-0.82	0.95	-1.45	-0.79	-0.16	0.34	1.28
1-2 years	664	0.17	-0.61	1.00	-1.29	-0.58	0.07	0.63	1.36	Postsecondary	1,865	0.49	-0.25	0.93	-0.84	-0.21	0.38	0.90	1.22
=>3 years	2,937	0.77	-0.41	0.97	-1.04	-0.36	0.26	0.80	1.29	≥University	390	0.10	0.05	0.98	-0.49	0.16	0.70	1.22	1.19
<b>Belgium Fr. 2011</b>	3,086	1.00	-0.47	0.95	-1.06	-0.43	0.18	0.71	1.25	<b>Belgium Fr. 2006</b>	3,086	1.00	-0.47	0.95	-1.06	-0.43	0.18	0.71	1.25
<=1 year	120	0.04	-0.94	1.06	-1.58	-0.96	-0.20	0.49	1.38	≤Secondary	1,083	0.35	-0.85	0.89	-1.41	-0.80	-0.23	0.26	1.17
1-2 years	583	0.19	-0.62	0.93	-1.20	-0.57	0.04	0.56	1.24	Postsecondary	393	0.13	-0.79	0.90	-1.35	-0.75	-0.19	0.37	1.16
=>3 years	2,383	0.77	-0.41	0.94	-1.00	-0.37	0.23	0.75	1.23	≥University	1,610	0.52	-0.12	0.86	-0.65	-0.09	0.47	0.93	1.12
<b>Belgium Fr. 2016</b>	3,666	1.00	-0.51	0.96	-1.13	-0.47	0.14	0.70	1.28	<b>Belgium Fr. 2006</b>	3,666	1.00	-0.51	0.96	-1.13	-0.47	0.14	0.70	1.28
<=1 year	66	0.02	-0.96	0.86	-1.54	-0.96	-0.39	-0.10	1.14	≤Secondary	1,070	0.29	-0.92	0.95	-1.51	-0.92	-0.28	0.29	1.23
1-2 years	187	0.05	-0.81	0.99	-1.51	-0.73	-0.10	0.40	1.40	Postsecondary	452	0.12	-0.89	0.84	-1.44	-0.87	-0.31	0.19	1.13
=>3 years	3,413	0.93	-0.49	0.96	-1.11	-0.45	0.17	0.72	1.27	≥University	2,144	0.58	-0.20	0.88	-0.75	-0.19	0.39	0.91	1.14
<b>France 2006</b>	3,733	1.00	-0.12	0.96	-0.74	-0.07	0.55	1.08	1.29	<b>France 2006</b>	3,763	1.00	-0.12	0.96	-0.74	-0.07	0.55	1.08	1.29
<=1 year	160	0.04	-0.22	1.01	-0.92	-0.24	0.52	1.16	1.44	≤Secondary	2,138	0.57	-0.40	0.92	-0.97	-0.36	0.24	0.77	1.21
1-2 years	739	0.20	-0.26	0.95	-0.88	-0.22	0.38	0.94	1.26	Postsecondary	585	0.16	0.09	0.89	-0.41	0.15	0.69	1.19	1.10
=>3 years	2,864	0.77	-0.07	0.96	-0.69	-0.01	0.59	1.11	1.28	≥University	1,040	0.28	0.38	0.87	-0.13	0.45	0.97	1.43	1.11
<b>France 2011</b>	3,927	1.00	-0.27	0.99	-0.93	-0.21	0.42	0.96	1.34	<b>France 2011</b>	3,763	1.00	-0.12	0.96	-0.74	-0.07	0.55	1.08	1.29
<=1 year	151	0.04	-0.39	1.06	-1.08	-0.35	0.35	0.88	1.43	≤Secondary	2,138	0.57	-0.40	0.92	-0.97	-0.36	0.24	0.77	1.21
1-2 years	807	0.21	-0.37	0.96	-1.01	-0.33	0.30	0.83	1.31	Postsecondary	585	0.16	0.09	0.89	-0.41	0.15	0.69	1.19	1.10
=>3 years	2,969	0.76	-0.24	1.00	-0.89	-0.18	0.45	1.00	1.34	≥University	1,040	0.28	0.38	0.87	-0.13	0.45	0.97	1.43	1.11
<b>France 2016</b>	4,061	1.00	-0.31	0.96	-0.91	-0.25	0.36	0.86	1.27	<b>France 2016</b>	4,061	1.00	-0.31	0.96	-0.91	-0.25	0.36	0.86	1.27
<=1 year	532	0.13	-0.62	0.99	-1.29	-0.55	0.08	0.58	1.37	≤Secondary	1,793	0.44	-0.64	0.94	-1.23	-0.58	0.00	0.50	1.23
1-2 years	2,579	0.64	-0.32	0.94	-0.91	-0.25	0.34	0.83	1.25	Postsecondary	815	0.20	-0.24	0.88	-0.79	-0.20	0.37	0.88	1.16
=>3 years	950	0.23	-0.14	0.96	-0.73	-0.08	0.52	1.02	1.25	≥University	1,453	0.36	0.04	0.91	-0.52	0.12	0.66	1.11	1.18

Notes and Sources: See end of Table A2.



**Table A2: End**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25
Preschool duration in years										Parental education levels									
<b>Ontario 2006</b>	3,346	1.00	0.35	1.02	-0.30	0.39	1.05	1.65	1.35	<b>Ontario 2006</b>	3,346	1.00	0.35	1.02	-0.30	0.39	1.05	1.65	1.35
<=1 year	2,763	0.83	0.34	1.02	-0.31	0.36	1.03	1.64	1.34	≤Secondary	607	0.18	0.03	1.03	-0.63	0.05	0.73	1.28	1.36
1-2 years	262	0.08	0.53	1.06	-0.77	0.54	1.19	1.79	1.96	Postsecondary	1,122	0.34	0.29	0.97	-0.30	0.32	0.94	1.56	1.24
=>3 years	321	0.10	0.39	1.03	-0.32	0.49	1.15	1.67	1.47	≥University	1,517	0.45	0.58	0.99	-0.08	0.61	1.28	1.84	1.35
<b>Ontario 2011</b>	3,478	1.00	0.24	1.03	-0.40	0.30	0.95	1.51	1.34	<b>Ontario 2011</b>	3,478	1.00	0.24	1.03	-0.40	0.30	0.95	1.51	1.34
<=1 year	2,585	0.74	0.17	1.04	-0.47	0.251	0.89	1.45	1.36	≤Secondary	484	0.14	-0.14	1.02	-0.83	-0.08	0.59	1.13	1.42
1-2 years	441	0.13	0.27	1.00	-0.36	0.26	0.96	1.51	1.33	Postsecondary	1,173	0.34	0.04	1.01	-0.55	0.15	0.70	1.22	1.25
=>3 years	652	0.19	0.42	1.01	-0.2	0.50	1.13	1.63	1.33	≥University	1,821	0.52	0.51	0.97	-0.12	0.55	1.20	1.73	1.32
<b>Ontario 2016</b>	3,287	1.00	0.19	1.08	-0.48	0.29	0.94	1.49	1.42	<b>Ontario 2016</b>	3,287	1.00	0.19	-0.48	0.29	0.94	1.49	1.47	1.20
<=1 year	859	0.26	0.00	1.08	-0.68	0.021	0.75	1.33	1.43	≤Secondary	326	0.10	-0.23	1.08	-0.96	-0.16	0.56	1.09	1.52
1-2 years	1,791	0.54	0.25	1.08	-0.41	0.361	1.00	1.55	1.41	Postsecondary	1,111	0.34	0.45	1.02	-0.19	0.53	1.15	1.72	1.34
=>3 years	637	0.19	0.28	1.04	-0.32	0.413	0.99	1.50	1.30	≥University	1,850	0.56	0.19	1.08	-0.48	0.29	0.94	1.49	1.42
<b>Québec 2006</b>	3,045	1.00	0.03	0.92	-0.55	0.07	0.67	1.17	1.22	<b>Québec 2006</b>	3,045	1.00	0.03	0.92	-0.55	0.07	0.67	1.17	1.22
<=1 year	2,463	0.81	0.03	0.92	-0.54	0.08	0.66	1.16	1.21	≤Secondary	649	0.21	-0.34	0.92	-0.94	-0.28	0.30	0.81	1.24
1-2 years	233	0.08	0.12	0.94	-0.51	0.08	0.84	1.34	1.36	Postsecondary	1,141	0.37	-0.08	0.88	-0.62	-0.04	0.52	1.01	1.14
=>3 years	349	0.11	-0.01	0.94	-0.60	0.02	0.64	1.10	1.23	≥University	1,255	0.41	0.31	0.88	-0.26	0.34	0.91	1.40	1.17
<b>Québec 2011</b>	3,603	1.00	-0.01	0.90	-0.59	0.02	0.60	1.10	1.19	<b>Québec 2011</b>	3,603	1.00	-0.01	0.90	-0.59	0.02	0.60	1.10	1.19
<=1 year	2,752	0.76	-0.05	0.89	-0.62	-0.02	0.55	1.06	1.18	≤Secondary	509	0.14	-0.40	0.86	-0.88	-0.35	0.19	0.63	1.07
1-2 years	304	0.08	0.12	0.86	-0.46	0.16	0.71	1.19	1.17	Postsecondary	1,396	0.39	-0.16	0.87	-0.72	-0.13	0.42	0.92	1.14
=>3 years	547	0.15	0.20	0.93	-0.39	0.21	0.83	1.34	1.23	≥University	1,698	0.47	0.25	0.86	-0.32	0.29	0.83	1.31	1.15
<b>Québec 2016</b>	2,687	1.00	0.19	0.90	-0.39	0.22	0.79	1.33	1.18	<b>Québec 2016</b>	2,687	1.00	0.19	0.90	-0.39	0.22	0.79	1.33	1.18
<=1 year	481	0.18	0.08	0.93	-0.48	0.09	0.71	1.24	1.19	≤Secondary	246	0.09	-0.28	0.94	-0.84	-0.23	0.38	0.91	1.23
1-2 years	1,040	0.39	0.17	0.91	-0.42	0.21	0.78	1.31	1.19	Postsecondary	978	0.36	0.04	0.84	-0.54	0.04	0.61	1.11	1.16
=>3 years	1,166	0.43	0.25	0.87	-0.34	0.28	0.84	1.38	1.18	≥University	1,463	0.54	0.40	0.87	-0.16	0.44	0.97	1.49	1.14

Notes: International z-scores are computed with house weight; percentile based on 5 plausible values and if student's information on parental education and preschool attendance are not missing. N: number of students in Grade 4 (and Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc.; P75-P25 the z-score gap between the P75 and P25 percentiles. The jurisdictions that did not participate in a year-survey are not indicated (see Table A). Attendance levels: Lower: 0 to <=1 years; Middle: 1-3 years; High: =>3 years. Parental education levels: Lower: ≤Secondary; Middle: Postsecondary; High: ≥University.

Sources: Authors' computation from weighted PIRLS 2006, 2011, 2016 data sets.

**Table A3: Percentile distribution of students' overall international math and science z-scores by jurisdiction and preschool duration levels, TIMSS 2015 and 2019**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Mean	SD	P25	P50	P75	P90	P75-P25
	Math									Science						
<b>Denmark 2015</b>	3,218	1.00	0.20	1.05	-0.48	0.25	0.93	1.52	1.41	0.00	0.97	-0.61	0.05	0.65	1.18	1.27
<=1 year	235	0.07	-0.06	1.08	-0.78	-0.06	0.63	1.33	1.41	-0.29	0.98	-0.86	-0.27	0.31	0.87	1.17
1-2 years	1,243	0.39	0.13	1.05	-0.53	0.19	0.86	1.42	1.39	-0.08	0.96	-0.70	-0.02	0.58	1.08	1.28
=>3 years	1,740	0.54	0.30	1.04	-0.40	0.33	1.01	1.60	1.41	0.10	0.95	-0.50	0.15	0.75	1.28	1.26
<b>Denmark 2019</b>	2,644	1.00	0.01	1.00	-0.66	0.03	0.70	1.28	1.36	0.01	0.99	-0.59	0.05	0.71	1.25	1.30
<=1 year	37	0.01	-0.49	1.29	-1.28	-0.26	0.17	1.40	1.45	-0.77	1.47	-1.33	-0.92	0.36	0.94	1.68
1-2 years	98	0.04	-0.26	0.97	-0.91	-0.21	0.34	0.84	1.25	-0.34	1.13	-1.26	-0.10	0.29	1.05	1.55
=>3 years	2,509	0.95	0.02	0.99	-0.64	0.05	0.71	1.28	1.36	0.04	0.97	-0.57	0.07	0.71	1.25	1.28
<b>Finland 2015</b>	4,704	1.00	0.14	0.94	-0.46	0.19	0.79	1.32	1.26	0.37	0.91	-0.19	0.41	0.97	1.48	1.16
<=1 year	1,393	0.30	0.15	0.93	-0.45	0.20	0.81	1.32	1.26	0.40	0.90	-0.17	0.43	0.99	1.53	1.16
1-2 years	1,118	0.24	0.07	0.95	-0.53	0.10	0.71	1.27	1.24	0.31	0.93	-0.23	0.35	0.92	1.42	1.15
=>3 years	2,193	0.47	0.17	0.94	-0.44	0.23	0.82	1.34	1.26	0.38	0.90	-0.19	0.43	0.98	1.48	1.16
<b>Finland 2019</b>	4,051	1.00	0.11	1.00	-0.52	0.13	0.79	1.38	1.32	0.40	0.93	-0.17	0.45	1.04	1.53	1.21
<=1 year	601	0.15	0.11	1.03	-0.55	0.16	0.85	1.37	1.40	0.43	0.96	-0.14	0.47	1.14	1.64	1.28
1-2 years	402	0.10	0.15	0.99	-0.47	0.13	0.83	1.38	1.30	0.43	0.99	-0.18	0.49	1.11	1.62	1.29
=>3 years	3,048	0.75	0.10	0.99	-0.52	0.13	0.79	1.38	1.32	0.40	0.91	-0.17	0.44	1.01	1.49	1.18
<b>Norway 5 2015</b>	1,812	1.00	0.47	0.99	-0.16	0.51	1.14	1.63	1.30	0.29	0.88	-0.24	0.32	0.88	1.43	1.30
<=1 year	130	0.07	0.05	1.10	-0.61	0.04	0.80	1.42	1.41	-0.18	1.10	-0.93	-0.17	0.59	1.03	1.41
1-2 years	387	0.19	0.40	0.94	-0.18	0.39	1.01	1.57	1.19	0.30	0.86	-0.28	0.24	0.94	1.45	1.19
=>3 years	1,295	0.76	0.53	0.98	-0.08	0.57	1.21	1.69	1.28	0.34	0.85	-0.18	0.37	0.89	1.44	1.28
<b>Norway 5 2019</b>	3,473	1.00	0.00	1.00	-0.64	0.05	0.71	1.24	1.35	0.00	1.00	-0.61	0.08	0.72	1.22	1.33
<=1 year	95	0.03	-0.34	1.02	-0.97	-0.42	0.49	1.00	1.46	-0.74	1.13	-1.67	-0.49	0.12	0.45	1.79
1-2 years	88	0.03	-0.31	0.90	-0.85	-0.37	0.27	0.91	1.12	-0.43	1.11	-1.15	-0.27	0.39	0.92	1.54
=>3 years	3,290	0.95	0.02	1.00	-0.62	0.07	0.72	1.25	1.34	0.04	0.98	-0.58	0.10	0.74	1.23	1.32
<b>Sweden 2015</b>	6,622	1.00	-0.01	0.93	-0.61	0.03	0.63	1.14	1.24	0.28	0.99	-0.33	0.34	0.95	1.48	1.27
<=1 year	626	0.09	-0.35	0.99	-1.05	-0.38	0.31	0.90	1.37	-0.19	1.14	-1.02	-0.19	0.62	1.29	1.65
1-2 years	924	0.14	-0.06	0.97	-0.64	0.03	0.60	1.10	1.24	0.23	1.00	-0.38	0.34	0.93	1.45	1.31
=>3 years	5,072	0.77	0.04	0.91	-0.54	0.08	0.67	1.17	1.21	0.35	0.95	-0.23	0.40	0.98	1.51	1.21
<b>Sweden 2019</b>	4,286	1.00	0.02	1.00	-0.62	0.04	0.70	1.26	1.32	0.02	0.99	-0.57	0.09	0.69	1.25	1.26
<=1 year	220	0.05	-0.52	1.05	-1.19	-0.52	0.19	0.83	1.38	-0.73	1.11	-1.49	-0.74	0.12	0.59	1.61
1-2 years	641	0.15	-0.12	1.00	-0.75	-0.05	0.63	1.12	1.38	-0.10	1.01	-0.69	0.02	0.56	1.11	1.25
=>3 years	3,425	0.80	0.08	0.98	-0.56	0.10	0.75	1.32	1.32	0.10	0.95	-0.48	0.15	0.75	1.29	1.24

Notes and Sources: See end of Table A3.

**Table A3: End**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Mean	SD	P25	P50	P75	P90	P75-P25
	Math									Science						
<b>Belgium Fl. 2015</b>	4,827	1.00	0.30	0.86	-0.27	0.31	0.88	1.40	1.16	-0.21	0.86	-0.77	-0.17	0.39	0.87	1.16
<=1 year	629	0.13	-0.08	0.89	-0.67	-0.08	0.51	1.01	1.17	-0.69	0.93	-1.32	-0.69	-0.02	0.49	1.30
1-2 years	1,337	0.28	0.16	0.82	-0.40	0.17	0.73	1.22	1.13	-0.34	0.83	-0.90	-0.32	0.24	0.71	1.14
=>3 years	2,861	0.59	0.45	0.83	-0.11	0.45	1.01	1.52	1.12	-0.04	0.81	-0.56	-0.02	0.51	0.98	1.07
<b>Belgium Fl. 2019</b>	3,926	1.00	0.12	0.88	-0.47	0.13	0.73	1.25	1.20	-0.29	0.88	-0.84	-0.24	0.33	0.79	1.17
<=1 year+	149	0.04	-0.39	0.91	-0.88	-0.42	0.06	0.79	0.94	-0.95	0.98	-1.61	-1.03	-0.29	0.38	1.32
1-2 years	207	0.05	-0.20	0.92	-0.94	-0.16	0.46	1.04	1.40	-0.51	0.86	-1.04	-0.45	0.07	0.57	1.12
=>3 years	3,570	0.91	0.16	0.87	-0.43	0.17	0.76	1.27	1.20	-0.25	0.86	-0.80	-0.20	0.36	0.81	1.16
<b>France 2015</b>	4,061	1.00	-0.49	1.04	-1.18	-0.45	0.25	0.84	1.43	-0.53	1.01	-1.19	-0.48	0.18	0.73	1.37
<=1 year	562	0.14	-0.77	1.10	-1.51	-0.77	-0.01	0.64	1.50	-0.75	1.07	-1.49	-0.68	-0.02	0.57	1.47
1-2 years	2,372	0.58	-0.47	1.02	-1.17	-0.44	0.26	0.84	1.43	-0.52	0.98	-1.17	-0.49	0.17	0.70	1.34
=>3 years	1,127	0.28	-0.37	1.03	-1.02	-0.33	0.33	0.92	1.35	-0.43	1.03	-1.10	-0.37	0.28	0.86	1.38
<b>France 2019</b>	5,229	1.00	-0.02	1.02	-0.70	0.03	0.70	1.28	1.40	0.00	1.03	-0.67	0.06	0.73	1.28	1.39
<=1 year	348	0.07	-0.36	1.04	-1.12	-0.31	0.39	0.98	1.51	-0.37	1.12	-1.18	-0.31	0.44	1.01	1.62
1-2 years	257	0.05	-0.23	1.16	-1.14	-0.14	0.61	1.31	1.75	-0.15	1.15	-1.04	0.07	0.68	1.28	1.73
=>3 years	4,624	0.88	0.02	1.01	-0.65	0.06	0.72	1.30	1.37	0.04	1.01	-0.61	0.08	0.75	1.30	1.36
<b>Ontario 2015</b>	3,525	1.00	-0.12	1.02	-0.79	-0.09	0.56	1.16	1.35	0.09	1.00	-0.55	0.15	0.79	1.34	1.34
<=1 year	2,096	0.59	-0.19	1.02	-0.88	-0.17	0.49	1.09	1.37	0.03	1.00	-0.63	0.09	0.73	1.26	1.36
1-2 years	801	0.23	-0.04	1.02	-0.69	0.00	0.62	1.21	1.31	0.19	1.02	-0.46	0.24	0.88	1.46	1.33
=>3 years	628	0.18	0.02	0.99	-0.63	0.08	0.69	1.27	1.32	0.21	0.98	-0.40	0.24	0.90	1.41	1.31
<b>Ontario 2019</b>	2,236	1.00	-0.04	1.03	-0.76	0.00	0.67	1.24	1.43	0.10	1.00	-0.53	0.15	0.78	1.35	1.31
<=1 year	683	0.31	-0.15	1.08	-0.89	-0.11	0.64	1.16	1.53	0.04	1.03	-0.63	0.12	0.75	1.29	1.38
1-2 years	660	0.30	-0.08	1.00	-0.74	-0.07	0.61	1.33	1.36	0.05	0.99	-0.59	0.08	0.73	1.35	1.32
=>3 years	893	0.40	0.07	1.00	-0.66	0.19	0.71	1.28	1.37	0.19	0.98	-0.44	0.21	0.84	1.44	1.28
<b>Québec 2015</b>	2,355	1.00	0.18	0.94	-0.43	0.18	0.81	1.37	1.24	-0.01	0.92	-0.61	-0.01	0.61	1.15	1.22
<=1 year	912	0.39	0.07	0.93	-0.51	0.06	0.70	1.22	1.22	-0.09	0.91	-0.76	-0.10	0.58	1.07	1.34
1-2 years	558	0.24	0.11	0.94	-0.43	0.11	0.72	1.26	1.15	-0.02	0.96	-0.60	0.02	0.63	1.17	1.23
=>3 years	885	0.38	0.32	0.95	-0.31	0.35	0.98	1.51	1.29	0.06	0.91	-0.53	0.06	0.68	1.20	1.21
<b>Québec 2019</b>	2,131	1.00	0.12	0.88	-0.49	0.17	0.70	1.24	1.19	-0.01	0.87	-0.61	0.03	0.61	1.08	1.22
<=1 year	430	0.20	0.01	0.92	-0.59	0.04	0.63	1.24	1.22	-0.10	0.91	-0.75	-0.05	0.56	1.12	1.31
1-2 years	309	0.15	0.01	0.96	-0.61	-0.03	0.68	1.28	1.29	-0.05	0.90	-0.70	-0.01	0.57	1.08	1.27
=>3 years	1,392	0.65	0.17	0.84	-0.40	0.21	0.73	1.24	1.13	0.02	0.85	-0.56	0.07	0.63	1.08	1.19

Notes: International z-scores are computed with house weight; percentile based on 5 plausible values and if student's information on parental education and preschool attendance are not missing. N: number of students in Grade 4 (Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc.; P75-P25 the z-score gap between the P75 and P25 percentiles. The jurisdictions that did not participate in a year-survey are not indicated (see Table A). Attendance levels: Lower: 0 to <2 years; Middle: 2-3 years; High: =>3 years. Source: Authors' computation from weighted TIMSS data sets 2015 and 2019.

**Table A4: Percentile distribution of students' overall math and science z-scores by jurisdiction and parental education levels, TIMSS 2015 and 2019**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Mean	SD	P25	P50	P75	P90	P75-P25	
			Math								Science						
<b>Denmark 2015</b>	3,218	1.00	0.20	1.05	-0.48	0.25	0.93	1.52	1.41	0.00	0.97	-0.61	0.05	0.65	1.18	1.27	
≤Secondary	347	0.11	-0.14	1.06	-0.82	-0.10	0.56	1.18	1.38	-0.36	1.01	-0.96	-0.32	0.32	0.86	1.28	
Postsecondary	838	0.26	0.03	1.00	-0.61	0.11	0.69	1.26	1.29	-0.17	0.92	-0.80	-0.12	0.48	0.98	1.28	
≥University	2,033	0.63	0.34	1.05	-0.36	0.38	1.06	1.64	1.42	0.13	0.95	-0.45	0.18	0.76	1.29	1.22	
<b>Denmark 2019</b>	2644	1.00	0.01	1.00	-0.66	0.03	0.70	1.28	1.36	0.01	0.99	-0.59	0.05	0.71	1.25	1.30	
≤Secondary	244	0.09	-0.44	0.97	-1.15	-0.34	0.21	0.69	1.36	-0.46	1.03	-1.05	-0.34	0.19	0.87	1.23	
Postsecondary	443	0.17	-0.31	0.99	-0.91	-0.37	0.32	1.02	1.23	-0.37	0.99	-1.00	-0.33	0.24	0.85	1.23	
≥University	1957	0.74	0.14	0.97	-0.49	0.17	0.79	1.37	1.28	0.16	0.95	-0.43	0.19	0.81	1.34	1.24	
<b>Finland 2015</b>	4,704	1.00	0.14	0.94	-0.46	0.19	0.79	1.32	1.26	0.37	0.91	-0.19	0.41	0.97	1.48	1.16	
≤Secondary	1,270	0.27	-0.17	0.93	-0.78	-0.11	0.48	0.95	1.25	0.06	0.89	-0.49	0.10	0.66	1.16	1.15	
Postsecondary	994	0.21	0.03	0.88	-0.55	0.04	0.63	1.17	1.18	0.30	0.86	-0.25	0.31	0.86	1.38	1.11	
≥University	2,440	0.52	0.37	0.91	-0.2	0.41	1.01	1.47	1.21	0.57	0.88	0.03	0.62	1.16	1.64	1.12	
<b>Finland 2019</b>	4,051	1.00	0.11	1.00	-0.52	0.13	0.79	1.38	1.32	0.40	0.93	-0.17	0.45	1.04	1.53	1.21	
≤Secondary	983	0.24	-0.21	0.96	-0.83	-0.23	0.43	1.02	1.27	0.13	0.93	-0.48	0.19	0.81	1.27	1.28	
Postsecondary	562	0.14	-0.21	0.97	-0.75	-0.14	0.42	0.97	1.17	0.17	0.88	-0.38	0.21	0.77	1.22	1.15	
≥University	2,506	0.62	0.31	0.97	-0.30	0.36	0.99	1.53	1.29	0.57	0.90	0.04	0.60	1.21	1.68	1.16	
<b>Norway 5 2015</b>	1,812	1.00	0.47	0.99	-0.16	0.51	1.14	1.63	1.30	0.29	0.88	-0.24	0.32	0.88	1.43	1.13	
≤Secondary	226	0.12	-0.15	1.07	-0.83	-0.13	0.59	1.32	1.42	-0.25	1.05	-0.95	-0.22	0.45	1.16	1.40	
Postsecondary	466	0.26	0.25	0.92	-0.36	0.31	0.93	1.39	1.29	0.13	0.87	-0.38	0.16	0.67	1.22	1.05	
≥University	1,120	0.62	0.69	0.92	0.07	0.70	1.28	1.86	1.21	0.48	0.79	-0.07	0.50	1.00	1.49	1.07	
<b>Norway 5 2019</b>	3,473	1.00	0.00	1.00	-0.64	0.05	0.71	1.24	1.35	0.00	1.00	-0.61	0.08	0.72	1.22	1.33	
≤Secondary	396	0.11	-0.46	0.91	-0.99	-0.45	0.14	0.71	1.13	-0.58	1.07	-1.23	-0.52	0.19	0.70	1.42	
Postsecondary	659	0.19	-0.28	0.93	-0.94	-0.23	0.40	0.90	1.34	-0.29	0.92	-0.92	-0.22	0.35	0.84	1.27	
≥University	2,418	0.70	0.18	0.99	-0.42	0.24	0.88	1.41	1.31	0.21	0.94	-0.36	0.28	0.87	1.35	1.23	
<b>Sweden 2015</b>	6,622	1.00	-0.01	0.93	-0.61	0.03	0.63	1.14	1.24	0.28	0.99	-0.33	0.34	0.95	1.48	1.27	
≤Secondary	1,416	0.21	-0.45	0.96	-1.09	-0.40	0.22	0.74	1.31	-0.20	1.08	-0.88	-0.14	0.55	1.12	1.43	
Postsecondary	1,636	0.25	-0.12	0.86	-0.64	-0.07	0.46	0.93	1.10	0.21	0.89	-0.35	0.25	0.80	1.28	1.15	
≥University	3,570	0.54	0.24	0.87	-0.33	0.27	0.85	1.34	1.18	0.53	0.91	-0.04	0.56	1.17	1.66	1.21	
<b>Sweden 2019</b>	4,286	1.00	0.02	1.00	-0.62	0.04	0.70	1.26	1.32	0.02	0.99	-0.57	0.09	0.69	1.25	1.26	
≤Secondary	857	0.20	-0.44	0.99	-1.07	-0.43	0.24	0.78	1.31	-0.47	1.05	-1.16	-0.40	0.28	0.79	1.44	
Postsecondary	984	0.23	-0.21	0.93	-0.83	-0.12	0.45	0.94	1.28	-0.18	0.93	-0.76	-0.14	0.45	1.00	1.21	
≥University	2,445	0.57	0.28	0.94	-0.34	0.33	0.94	1.48	1.28	0.29	0.90	-0.27	0.31	0.92	1.43	1.19	

Notes and Sources: See end of Table A4.

**Table A4: End**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	P75-P25	Mean	SD	P25	P50	P75	P90	P75-P25	
			Math								Science						
<b>Belgium Fl. 2015</b>	4,827	1.00	0.30	0.86	-0.27	0.31	0.88	1.40	1.16	-0.21	0.86	-0.77	-0.17	0.39	0.87	1.16	
≤Secondary	1,176	0.24	-0.12	0.79	-0.64	-0.11	0.41	0.89	1.05	-0.67	0.82	-1.22	-0.66	-0.09	0.38	1.13	
Postsecondary	968	0.20	0.09	0.77	-0.42	0.10	0.63	1.06	1.05	-0.41	0.77	-0.94	-0.38	0.13	0.54	1.07	
≥University	2,683	0.56	0.55	0.82	0.01	0.57	1.11	1.58	1.10	0.06	0.81	-0.44	0.09	0.61	1.06	1.06	
<b>Belgium Fl. 2019</b>	3,926	1.00	0.12	0.88	-0.47	0.13	0.73	1.25	1.20	-0.29	0.88	-0.84	-0.24	0.33	0.79	1.17	
≤Secondary	825	0.21	-0.29	0.86	-0.85	-0.25	0.26	0.77	1.11	-0.74	0.90	-1.36	-0.74	-0.12	0.45	1.24	
Postsecondary	712	0.18	-0.17	0.81	-0.74	-0.21	0.39	0.87	1.13	-0.55	0.82	-1.07	-0.53	-0.02	0.51	1.05	
≥University	2,389	0.61	0.35	0.83	-0.19	0.38	0.92	1.39	1.11	-0.05	0.80	-0.56	-0.02	0.49	0.91	1.05	
<b>France 2015</b>	4,061	1.00	-0.49	1.04	-1.18	-0.45	0.25	0.84	1.43	-0.53	1.01	-1.19	-0.48	0.18	0.73	1.37	
≤Secondary	1,869	0.46	-0.90	1.01	-1.58	-0.87	-0.21	0.39	1.37	-0.92	0.98	-1.58	-0.90	-0.24	0.33	1.34	
Postsecondary	853	0.21	-0.41	0.94	-1.03	-0.39	0.25	0.77	1.28	-0.49	0.91	-1.04	-0.44	0.16	0.61	1.20	
≥University	1,339	0.33	0.02	0.91	-0.57	0.07	0.67	1.16	1.24	-0.04	0.89	-0.61	0.00	0.57	1.06	1.18	
<b>France 2019</b>	5,229	1.00	-0.02	1.02	-0.70	0.03	0.70	1.28	1.40	0.00	1.03	-0.67	0.06	0.73	1.28	1.39	
≤Secondary	1,992	0.38	-0.49	0.97	-1.17	-0.48	0.17	0.73	1.35	-0.49	0.99	-1.15	-0.46	0.19	0.75	1.33	
Postsecondary	1,096	0.21	0.02	0.90	-0.60	0.08	0.68	1.12	1.28	0.05	0.91	-0.54	0.08	0.69	1.19	1.23	
≥University	2,141	0.41	0.48	0.91	-0.09	0.52	1.12	1.61	1.21	0.51	0.87	-0.02	0.55	1.11	1.59	1.13	
<b>Ontario 2015</b>	3,525	1.00	-0.12	1.02	-0.79	-0.09	0.56	1.16	1.35	0.09	1.00	-0.55	0.15	0.79	1.34	1.34	
≤Secondary	459	0.13	-0.52	1.02	-1.19	-0.49	0.18	0.76	1.37	-0.27	1.04	-0.94	-0.24	0.47	1.03	1.42	
Postsecondary	1,153	0.33	-0.34	0.97	-1.00	-0.3	0.34	0.88	1.34	-0.09	0.97	-0.72	-0.06	0.60	1.14	1.31	
≥University	1,913	0.54	0.11	0.98	-0.52	0.15	0.77	1.35	1.29	0.30	0.96	-0.31	0.36	0.95	1.49	1.26	
<b>Ontario 2019</b>	2,236	1.00	-0.04	1.03	-0.76	0.00	0.67	1.24	1.43	0.10	1.00	-0.53	0.15	0.78	1.35	1.31	
≤Secondary	188	0.08	-0.54	1.04	-1.28	-0.47	0.21	0.70	1.49	-0.39	0.98	-1.10	-0.39	0.38	0.76	1.48	
Postsecondary	668	0.30	-0.39	0.96	-0.98	-0.42	0.29	0.80	1.26	-0.11	0.98	-0.74	-0.07	0.53	1.07	1.27	
≥University	1,380	0.62	0.21	0.99	-0.45	0.31	0.86	1.45	1.31	0.29	0.97	-0.31	0.33	0.99	1.49	1.30	
<b>Québec 2015</b>	2,355	1.00	0.18	0.94	-0.43	0.18	0.81	1.37	1.24	-0.01	0.92	-0.61	-0.01	0.61	1.15	1.22	
≤Secondary	246	0.10	-0.32	0.94	-0.91	-0.29	0.32	0.85	1.22	-0.52	0.91	-1.10	-0.47	0.06	0.60	1.16	
Postsecondary	767	0.33	-0.02	0.86	-0.55	-0.03	0.56	1.07	1.11	-0.19	0.84	-0.75	-0.17	0.38	0.84	1.13	
≥University	1,342	0.57	0.40	0.92	-0.23	0.42	1.03	1.57	1.26	0.20	0.90	-0.39	0.21	0.83	1.33	1.22	
<b>Québec 2019</b>	2,196	1.00	0.11	0.88	-0.49	0.17	0.70	1.22	1.19	-0.02	0.88	-0.61	0.03	0.61	1.07	1.22	
≤Secondary	196	0.09	-0.53	0.87	-1.10	-0.53	0.16	0.63	1.25	-0.54	0.81	-1.10	-0.56	0.01	0.47	1.11	
Postsecondary	728	0.33	-0.08	0.81	-0.64	-0.06	0.48	0.98	1.12	-0.16	0.85	-0.76	-0.08	0.44	0.87	1.20	
≥University	1,272	0.58	0.35	0.84	-0.23	0.36	0.91	1.44	1.14	0.17	0.85	-0.38	0.19	0.76	1.28	1.15	

Notes: International z-scores are computed with house weight; percentile based on 5 plausible values and if student's information on parental education and preschool attendance are not missing. N: number of students in Grade 4 (Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc.; P75-P25 the z-score gap between the P75 and P25 percentiles.

Sources: Authors' computation from weighted TIMSS data sets 2015 and 2019.

**Table A5: Percentile distribution of students' overall international reading z-scores by jurisdiction and tercile of parental socioeconomic levels, PIRLS 2006-2011-2016**

Statistics	N	%N	Mean	SD	P25	P50	P75	P90	N	%N	Mean	SD	P25	P50	P75	P90	N	%N	Mean	SD	P25	P50	P75	P90
Jurisdiction	Terciles of socioeconomic status 2006								Terciles of socioeconomic status 2011								Terciles of socioeconomic status 2016							
<b>Denmark</b>	3,248	1	0.24	1.00	-0.34	0.35	0.90	1.43	3,823	1	0.24	0.94	-0.35	0.31	0.91	1.38	2,924	1	0.17	0.94	-0.40	0.25	0.79	1.31
Tercile 1	1,015	0.31	-0.04	1.00	-0.66	0.08	0.67	1.11	1,255	0.33	-0.07	0.94	-0.70	0.02	0.61	1.04	957	0.33	-0.12	0.94	-0.74	-0.05	0.52	0.97
Tercile 2	1,049	0.32	0.24	1.00	-0.29	0.32	0.91	1.43	1,255	0.33	0.20	0.91	-0.38	0.27	0.82	1.26	996	0.34	0.15	0.92	-0.39	0.26	0.74	1.20
Tercile 3	1,184	0.36	0.52	0.93	-0.04	0.61	1.10	1.69	1,313	0.34	0.60	0.86	0.05	0.64	1.19	1.67	971	0.33	0.49	0.87	-0.05	0.51	1.10	1.55
<b>Finland</b>	NA								4,387	1	0.41	0.93	-0.18	0.46	1.07	1.57	4,408	1	0.43	0.90	-0.13	0.49	1.04	1.53
Tercile 1									1,355	0.31	0.22	0.90	-0.32	0.26	0.81	1.31	1,425	0.32	0.19	0.93	-0.39	0.24	0.85	1.35
Tercile 2									1,436	0.33	0.35	0.95	-0.30	0.41	1.01	1.52	1,462	0.33	0.49	0.87	-0.11	0.57	1.08	1.53
Tercile 3									1,596	0.36	0.67	0.90	0.12	0.71	1.31	1.78	1,521	0.35	0.62	0.85	0.09	0.64	1.17	1.69
<b>Norway 4#/5</b>	3,264#	1	-0.48	0.95	-1.04	-0.41	0.16	0.67	2,696#	1	-0.46	0.90	-1.06	-0.41	0.19	0.63	3,623	1	0.32	0.93	-0.27	0.37	0.94	1.46
Tercile 1	1,072	0.33	-0.73	0.97	-1.33	-0.62	-0.07	0.41	846	0.31	-0.67	0.91	-1.28	-0.65	-0.08	0.47	1,148	0.32	0.04	0.96	-0.55	0.12	0.72	1.11
Tercile 2	1,128	0.35	-0.46	0.94	-1.01	-0.38	0.17	0.65	917	0.34	-0.40	0.90	-1.05	-0.39	0.25	0.70	1,245	0.34	0.43	0.89	-0.15	0.43	1.00	1.54
Tercile 3	1,064	0.33	-0.23	0.89	-0.78	-0.19	0.35	0.83	933	0.35	-0.29	0.84	-0.81	-0.27	0.29	0.73	1,230	0.34	0.50	0.88	-0.07	0.54	1.11	1.59
<b>Sweden</b>	3,585	1	0.28	0.92	-0.27	0.33	0.91	1.40	3,523	1	0.10	0.93	-0.45	0.15	0.72	1.21	3,491	1	0.31	0.92	-0.23	0.39	0.94	1.43
Tercile 1	1,155	0.32	0.02	0.89	-0.48	0.07	0.60	1.14	1,163	0.33	-0.16	0.91	-0.76	-0.13	0.45	0.98	1,170	0.34	0.00	0.91	-0.52	0.08	0.61	1.12
Tercile 2	1,184	0.33	0.23	0.94	-0.38	0.28	0.88	1.41	1,190	0.34	0.08	0.92	-0.45	0.15	0.68	1.19	1,149	0.33	0.38	0.93	-0.18	0.47	1.01	1.44
Tercile 3	1,246	0.35	0.58	0.84	0.07	0.60	1.17	1.60	1,170	0.33	0.39	0.87	-0.15	0.38	0.98	1.45	1,172	0.34	0.56	0.84	0.02	0.61	1.13	1.60
<b>Belgium Fl.</b>	3,961	1	0.24	0.79	-0.28	0.26	0.78	1.22	NA								4,313	1	-0.13	0.84	-0.67	-0.09	0.45	0.91
Tercile 1	1,304	0.33	-0.04	0.79	-0.56	-0.04	0.47	0.96									1,434	0.33	-0.37	0.83	-0.89	-0.32	0.21	0.62
Tercile 2	1,328	0.34	0.27	0.75	-0.17	0.28	0.76	1.22									1,433	0.33	-0.15	0.80	-0.68	-0.14	0.42	0.87
Tercile 3	1,329	0.34	0.49	0.73	0.00	0.53	0.99	1.39									1,446	0.34	0.13	0.80	-0.38	0.16	0.71	1.12
<b>Belgium Fr.</b>	3,507	1	-0.41	0.99	-1.03	-0.37	0.27	0.83	2,911	1	-0.46	0.94	-1.04	-0.43	0.18	0.73	3,676	1	-0.52	0.97	-1.15	-0.48	0.14	0.71
Tercile 1	1,120	0.32	-0.83	0.95	-1.46	-0.81	-0.15	0.31	925	0.32	-0.88	0.90	-1.45	-0.85	-0.28	0.28	1,179	0.32	-0.93	0.92	-1.54	-0.93	-0.28	0.22
Tercile 2	1,179	0.34	-0.42	0.92	-1.00	-0.39	0.22	0.75	1,003	0.34	-0.46	0.89	-1.01	-0.45	0.13	0.69	1,254	0.34	-0.51	0.93	-1.07	-0.42	0.11	0.65
Tercile 3	1,208	0.34	0.03	0.91	-0.52	0.09	0.65	1.13	983	0.34	-0.03	0.82	-0.55	0.02	0.55	0.99	1,243	0.34	-0.12	0.87	-0.68	-0.14	0.48	0.99
<b>France</b>	3,456	1	-0.08	0.96	-0.69	-0.04	0.58	1.10	3,589	1	-0.21	0.97	-0.84	-0.14	0.44	0.99	3,908	1	-0.33	0.95	-0.93	-0.26	0.33	0.84
Tercile 1	1,102	0.32	-0.50	0.93	-1.07	-0.50	0.12	0.62	1,136	0.32	-0.61	0.93	-1.23	-0.59	0.03	0.47	1,346	0.34	-0.65	0.93	-1.27	-0.59	-0.01	0.51
Tercile 2	1,153	0.33	-0.06	0.89	-0.62	-0.06	0.56	1.03	1,218	0.34	-0.22	0.91	-0.79	-0.16	0.41	0.90	1,326	0.34	-0.33	0.91	-0.92	-0.27	0.31	0.83
Tercile 3	1,201	0.35	0.32	0.88	-0.20	0.39	0.93	1.39	1,235	0.34	0.20	0.89	-0.39	0.26	0.82	1.33	1,236	0.32	-0.02	0.90	-0.57	0.08	0.57	1.04
<b>Ontario</b>	3,264	1	0.37	1.01	-0.30	0.40	1.05	1.64	3,344	1	0.24	1.02	-0.40	0.34	0.94	1.47	3,097	1	0.19	1.07	-0.49	0.28	0.91	1.51
Tercile 1	1,025	0.31	0.13	0.96	-0.47	0.20	0.77	1.31	1,065	0.32	-0.04	1.04	-0.67	0.03	0.66	1.25	950	0.31	-0.04	1.07	-0.74	0.08	0.70	1.22
Tercile 2	990	0.30	0.41	1.04	-0.21	0.44	1.12	1.74	1,070	0.32	0.28	0.98	-0.33	0.38	0.96	1.48	994	0.32	0.21	1.06	-0.42	0.28	0.97	1.50
Tercile 3	1,249	0.38	0.56	0.97	-0.10	0.62	1.28	1.75	1,209	0.36	0.47	0.98	-0.13	0.55	1.15	1.70	1,153	0.37	0.38	1.05	-0.35	0.48	1.10	1.70
<b>Québec</b>	2,928	1	0.05	0.93	-0.54	0.11	0.68	1.21	3,429	1	0.00	0.90	-0.57	0.05	0.61	1.11	2,544	1	0.18	0.88	-0.40	0.23	0.78	1.32
Tercile 1	929	0.32	-0.19	0.90	-0.75	-0.10	0.44	0.93	1,065	0.31	-0.21	0.88	-0.74	-0.16	0.40	0.84	776	0.31	0.00	0.90	-0.57	-0.01	0.66	1.20
Tercile 2	1,089	0.37	0.02	0.91	-0.55	0.08	0.66	1.21	1,162	0.34	-0.04	0.88	-0.64	-0.04	0.54	1.06	921	0.36	0.21	0.84	-0.33	0.25	0.74	1.27
Tercile 3	910	0.31	0.32	0.91	-0.29	0.34	0.93	1.52	1,202	0.35	0.27	0.87	-0.28	0.32	0.84	1.32	847	0.33	0.34	0.87	-0.25	0.37	0.90	1.43

Notes: International z-scores are computed with house weight; percentile based on 5 plausible values and if student's information on parental occupation are not missing. N: number of students in Grade 4 (Grade4# and Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc. NA: Indicates the jurisdiction did not participate in this survey (see Table A).

Sources: Authors' computation from weighted PIRLS data sets 2006, 2011, and 2016.

**Table A6: Percentile distribution of students' overall international math and science z-scores by jurisdiction and terciles of parental socioeconomic status, TIMSS 2015 and 2019**

Statistics	N	%N	Math							Science						
			Mean	SD	P25	P50	P75	P90	P75-P25	Mean	SD	P25	P50	P75	P90	P75-P25
<b>Denmark</b>																
<b>2015</b>	3,132	1	0.19	1.05	-0.48	0.21	0.92	1.47	1.41	-0.02	0.98	-0.64	0.02	0.65	1.17	1.35
Tercile 1	1,035	0.33	-0.10	1.02	-0.76	-0.08	0.60	1.21	1.36	-0.32	0.99	-0.98	-0.27	0.38	0.95	1.40
Tercile 2	1,036	0.33	0.15	0.99	-0.46	0.19	0.82	1.39	1.27	-0.02	0.92	-0.62	-0.01	0.55	1.10	1.32
Tercile 3	1,061	0.34	0.51	1.03	-0.11	0.57	1.23	1.76	1.33	0.29	0.94	-0.29	0.32	0.93	1.45	1.27
<b>2019</b>	2,683	1.00	0.01	0.99	-0.65	0.05	0.71	1.28	1.36	0.01	0.99	-0.65	0.05	0.71	1.28	1.36
Tercile 1	877	0.33	-0.27	0.99	-0.92	-0.28	0.43	0.92	1.35	-0.27	0.99	-0.92	-0.28	0.43	0.92	1.35
Tercile 2	878	0.33	0.04	0.96	-0.60	0.04	0.73	1.31	1.33	0.04	0.96	-0.60	0.04	0.73	1.31	1.33
Tercile 3	928	0.35	0.27	0.96	-0.34	0.30	0.90	1.52	1.23	0.27	0.96	-0.34	0.30	0.90	1.52	1.23
<b>Finland</b>																
<b>2015</b>	4,588	1.00	0.16	0.93	-0.44	0.20	0.80	1.33	1.25	0.37	0.90	-0.19	0.41	0.97	1.50	1.16
Tercile 1	1,509	0.33	-0.04	0.95	-0.63	0.02	0.61	1.12	1.24	0.16	0.91	-0.40	0.17	0.77	1.25	1.16
Tercile 2	1,025	0.22	0.16	0.92	-0.44	0.19	0.79	1.35	1.23	0.43	0.90	-0.14	0.50	1.05	1.55	1.19
Tercile 3	2,054	0.45	0.31	0.90	-0.29	0.355	0.95	1.42	1.24	0.52	0.84	-0.01	0.54	1.08	1.60	1.09
<b>2019</b>	3,992	1	0.12	0.99	-0.50	0.15	0.80	1.38	1.30	0.43	0.91	-0.13	0.47	1.05	1.54	1.18
Tercile 1	1,252	0.31	-0.11	0.97	-0.73	-0.12	0.55	1.14	1.28	0.21	0.92	-0.36	0.25	0.87	1.33	1.23
Tercile 2	1,347	0.34	0.09	0.98	-0.53	0.14	0.76	1.32	1.29	0.40	0.88	-0.15	0.47	1.01	1.46	1.16
Tercile 3	1,393	0.35	0.40	0.95	-0.18	0.43	1.06	1.58	1.24	0.66	0.88	0.15	0.66	1.29	1.79	1.13
<b>Norway 5</b>																
<b>2015</b>	1,754	1	0.49	0.98	-0.13	0.53	1.15	1.65	1.28	0.30	0.88	-0.23	0.35	0.90	1.44	1.13
Tercile 1	552	0.31	0.16	0.97	-0.48	0.11	0.85	1.39	1.34	0.01	0.93	-0.59	0.03	0.64	1.21	1.23
Tercile 2	618	0.35	0.65	0.97	0.09	0.71	1.26	1.82	1.16	0.45	0.84	-0.07	0.45	0.99	1.52	1.06
Tercile 3	584	0.33	0.66	0.93	0.05	0.67	1.28	1.74	1.23	0.46	0.79	-0.09	0.48	0.99	1.44	1.08
<b>Norway 5</b>																
<b>2019</b>	NA									NA						
<b>Sweden</b>																
<b>2015</b>	6,412	1	0.00	0.91	-0.59	0.03	0.62	1.14	1.21	0.32	0.96	-0.27	0.38	0.97	1.51	1.24
Tercile 1	2,026	0.32	-0.28	0.92	-0.90	-0.20	0.34	0.84	1.24	0.05	1.00	-0.55	0.13	0.71	1.23	1.26
Tercile 2	2,161	0.34	0.06	0.88	-0.49	0.12	0.69	1.12	1.18	0.34	0.90	-0.23	0.42	0.96	1.47	1.19
Tercile 3	2,225	0.35	0.20	0.88	-0.41	0.21	0.86	1.34	1.27	0.56	0.92	-0.04	0.58	1.20	1.72	1.23
<b>2019</b>	4,196	1.00	0.03	0.99	-0.60	0.05	0.72	1.27	1.32	0.04	0.98	-0.55	0.10	0.71	1.26	1.26
Tercile 1	1,325	0.32	-0.29	1.01	-0.96	-0.23	0.42	0.94	1.38	-0.28	1.02	-0.95	-0.17	0.41	0.97	1.36
Tercile 2	1,401	0.33	0.05	0.95	-0.58	0.04	0.71	1.24	1.29	0.05	0.94	-0.51	0.09	0.69	1.24	1.21
Tercile 3	1,470	0.35	0.33	0.91	-0.27	0.38	0.97	1.50	1.24	0.35	0.87	-0.18	0.37	0.98	1.44	1.16
<b>Belgium Fl.</b>																
<b>2015</b>	4,576	1	0.30	0.84	-0.25	0.32	0.87	1.39	1.12	-0.18	0.84	-0.74	-0.14	0.39	0.86	1.13
Tercile 1	1,530	0.33	-0.02	0.80	-0.55	-0.02	0.53	0.98	1.09	-0.55	0.82	-1.10	-0.51	0.03	0.45	1.12
Tercile 2	1,523	0.33	0.36	0.79	-0.15	0.39	0.89	1.38	1.04	-0.12	0.78	-0.61	-0.09	0.41	0.87	1.02
Tercile 3	1,523	0.33	0.56	0.81	0.05	0.56	1.11	1.62	1.06	0.11	0.79	-0.41	0.15	0.65	1.11	1.06
<b>2019</b>	3,899	1.00	0.13	0.88	-0.47	0.13	0.74	1.26	1.21	-0.28	0.88	-0.83	-0.23	0.34	0.80	1.17
Tercile 1	1,303	0.33	-0.21	0.84	-0.81	-0.19	0.39	0.84	1.20	-0.64	0.87	-1.24	-0.62	-0.03	0.47	1.21
Tercile 2	1,304	0.33	0.22	0.88	-0.38	0.20	0.86	1.35	1.25	-0.19	0.84	-0.74	-0.14	0.41	0.82	1.15
Tercile 3	1,292	0.33	0.38	0.81	-0.11	0.40	0.93	1.38	1.04	-0.01	0.79	-0.47	0.01	0.53	0.97	1.00

Notes and Sources: See end of Table A6.

**Table A6: End**

Statistics	N	%N	Math						P75-	Science						P75-
			Mean	SD	P25	P50	P75	P90	P25	Mean	SD	P25	P50	P75	P90	P25
<b>France</b>																
<b>2015</b>	3,700	1	-0.47	1.02	-1.15	-0.43	0.26	0.83	1.41	-0.52	1.01	-1.17	-0.46	0.18	0.72	1.35
Tercile 1	1,254	0.34	-0.86	0.99	-1.50	-0.82	-0.18	0.43	1.32	-0.91	0.97	-1.56	-0.91	-0.18	0.29	1.38
Tercile 2	1,218	0.33	-0.50	0.95	-1.10	-0.46	0.14	0.69	1.24	-0.54	0.94	-1.15	-0.51	0.10	0.66	1.25
Tercile 3	1,228	0.33	-0.04	0.96	-0.69	0.02	0.65	1.09	1.34	-0.10	0.94	-0.67	-0.04	0.54	1.06	1.21
<b>2019</b>	4,986	1.00	0.00	1.02	-0.68	0.05	0.71	1.29	1.39	0.02	1.02	-0.66	0.07	0.73	1.29	1.39
Tercile 1	1,509	0.30	-0.41	1.00	-1.11	-0.39	0.27	0.81	1.37	-0.43	1.02	-1.06	-0.42	0.27	0.83	1.33
Tercile 2	1,634	0.33	0.02	0.92	-0.61	0.07	0.68	1.18	1.29	0.03	0.92	-0.57	0.10	0.69	1.18	1.26
Tercile 3	1,843	0.37	0.38	0.97	-0.24	0.47	1.07	1.58	1.31	0.42	0.95	-0.16	0.47	1.09	1.62	1.24
<b>Ontario</b>																
<b>2015</b>	3,427	1	-0.14	1.02	-0.81	-0.11	0.55	1.12	1.35	0.08	1.00	-0.57	0.13	0.77	1.32	1.34
Tercile 1	1,147	0.33	-0.42	1.00	-1.09	-0.42	0.31	0.84	1.40	-0.14	1.01	-0.81	-0.13	0.58	1.13	1.39
Tercile 2	1,128	0.33	-0.13	1.00	-0.78	-0.06	0.54	1.13	1.32	0.04	0.99	-0.59	0.12	0.74	1.25	1.32
Tercile 3	1,152	0.34	0.12	0.97	-0.51	0.15	0.77	1.36	1.27	0.34	0.94	-0.29	0.40	0.96	1.59	1.25
<b>2019</b>	2,408	1.00	-0.05	1.02	-0.76	-0.01	0.67	1.23	1.43	0.10	1.00	-0.54	0.15	0.78	1.35	1.31
Tercile 1	757	0.31	-0.28	1.06	-1.02	-0.25	0.46	1.10	1.48	-0.07	1.03	-0.74	-0.02	0.68	1.25	1.42
Tercile 2	790	0.33	0.00	0.97	-0.68	0.06	0.64	1.19	1.32	0.10	0.93	-0.51	0.20	0.70	1.23	1.21
Tercile 3	861	0.36	0.14	0.98	-0.53	0.19	0.78	1.33	1.31	0.28	1.00	-0.32	0.32	0.98	1.49	1.30
<b>Québec</b>																
<b>2015</b>	2,302	1	0.16	0.93	-0.47	0.16	0.79	1.32	1	-0.03	0.91	-0.62	0.00	0.58	1.16	1.35
Tercile 1	681	0.30	-0.12	0.87	-0.74	-0.12	0.52	0.99	0.75	-0.28	0.86	-0.86	-0.27	0.30	0.78	1.38
Tercile 2	857	0.37	0.11	0.95	-0.45	0.12	0.73	1.25	0.95	-0.07	0.93	-0.62	-0.05	0.51	1.14	1.25
Tercile 3	764	0.33	0.49	0.86	-0.13	0.50	1.14	1.48	1.12	0.27	0.85	-0.33	0.26	0.87	1.34	1.21
<b>2019</b>	2,701	1	0.13	0.88	-0.47	0.18	0.71	1.24	1.18	-0.01	0.88	-0.61	0.03	0.61	1.12	1.22
Tercile 1	836	0.31	-0.06	0.87	-0.66	-0.03	0.57	1.04	1.23	-0.17	0.85	-0.79	-0.14	0.43	0.88	1.22
Tercile 2	895	0.33	0.11	0.85	-0.46	0.17	0.67	1.19	1.13	-0.06	0.88	-0.65	-0.02	0.58	1.03	1.23
Tercile 3	970	0.36	0.34	0.86	-0.23	0.36	0.91	1.45	1.15	0.21	0.86	-0.34	0.23	0.81	1.37	1.15

Notes: International z-scores are computed with house weight; percentile based on 5 plausible values and if student's information on parental occupation and education are not missing. N: number of students in Grade 4 (Grade 5 in Norway) with test scores and parental covariates; SD: standard deviation; P10 indicates the 10th percentile of the distribution, P25 the 25th percentile, etc. NA: Indicates the jurisdiction did participate in this survey but the occupational informal is not available. Sources: Authors' computation from weighted TIMSS data sets 2015 and 2019.



## Appendix 1: Empirical studies on preschool programs effects for selected countries

### Country studies

In *England*, the introduction of free part-time places in a nursery or other registered setting for 3-year-olds are found to have small effects (Blanden, et al., 2014); while attendance in preschool moderately improves test results at ages 11, 14, and 16 (Apps, Mendolia, and Walker, 2013).

In the *Netherlands*, the expansion of preschool opportunities for children aged 4 generates a positive impact for disadvantaged children only (Leuven, Lindahl, Oosterbeek, and Webbink, 2010).

In *Spain*, Felfe, Nollenberger, and Rodriguez-Planas (2015) consider a sizeable expansion of publicly subsidized full-time high-quality childcare for 3-year-olds in the early 1990s. Using a difference-in-difference approach, they find evidence for improvements in PISA students' reading and math skills at age 15 from different cohorts and surveys. The effects are driven by girls and disadvantaged children. Santin and Cecilia (2018) use an identification strategy that reproduces a natural experiment on number of preschool year's impact for Grade 4 students. Comparing classrooms in a national 2009 schools survey, they find that the group with more average years of preschool education have significantly and positively higher average outcomes in reading and math. Moreover, preschool attendance before 3 years old is related with socio-economic variables.

In *Italy*, Carta and Rizzica (2018) analyze early kindergarten early access to subsidized childcare for 2-year-old children; results show no impacts on children's cognitive development, regardless of their family background. Aliverninia and Manganellia (2015), from PIRLS data on Grade 4 students, find that parental SES has a significant positive direct impact on reading literacy at both the student and school level, while pre-primary education does not have a significant impact on reading literacy.

In *France*, two studies using government panel surveys, providing information on children's cognitive and non-cognitive achievement when they enter first grade, were conducted to examine the effects of the extension of preschool to 2-year-olds (Dumas and Lefranc, 2012; Filatriau, Fougère, and Tô, 2013). They find that children who spend four rather than three years in preschool have higher cognitive and non-cognitive skills when entering first grade in elementary school. However, impacts did not differ by parental attributes or when the policy was mainly implemented to develop preschool spaces in a low-income neighborhood.

In *Germany*, several studies (Cornelissen et al., 2018; Felfe and Lalive, 2012, 2018) have explored the impact of the expansion of low fee childcare on child outcomes, in particular policy reforms of the childcare system entitling, since 1996, every child to a space in kindergarten from age 3 until primary school; and, in December 2008, all children aged 1 year or more to have a legal claim to a childcare slot by August 2013. Despite the large increase of childcare spaces with fees remaining constant, coverage remains low in Germany with aggregate demand still exceeding supply (as of year 2015), which is an issue to consider when interpreting treatment effects. The main framework adopted estimates marginal treatment effects (MTE)<sup>23</sup> which informs on their heterogeneity. The authors affirm their results point to a pattern of reverse selection on gains in terms of observed characteristics. Children, less likely to attend childcare than those from advantaged backgrounds (mostly children of Turkish descent, of single mothers or mothers with a low education level) have larger positive treatment effects because of their worse outcomes when not enrolled in childcare. Both studies suggest that the universal childcare program disproportionately subsidizes advantaged families whose children have the least to gain from early child care attendance. Children least likely to attend daycare early on, are from families with a higher resistance to enrolling a child, are found to benefit positively, with stronger effects for those with the highest resistance to treatment.

However, there are some drawbacks to this German policy with lessons for universal childcare policies that may be flimsy in other societies. First, the main outcome studied is *school readiness*, a mandatory school entry examination by pediatricians who assess children's development with regard to language skills, motor skills, and socio-emotional skills in the form of a medical diagnosis. The examination informs parents and schools about children's readiness to follow the curriculum taught in primary schools. This is a different outcome than results from more formal tests. Second, subsidized public childcare, considered as of good quality, offers generally

---

<sup>23</sup> The MTE approach requires specific data hard to obtain. Data sets used are not nation-wide but slots created for children ages 1 to 6 in some states.

spaces for half-day attendance (3-4 hours in the morning). Parents, the family or other types of care providers, must attend children for the rest of the day. This organization is quite different from the norms of policies offering full-day, full-year childcare, and public school pre-kindergarten.

In West German states using data from a socio-economic panel, Busse and Gathmann (2018), show that increases in childcare attendance for the youngest children (2-3-year-old) mirrored by the decline in exclusive care at home, display few substantive impacts on child development and mothers' labor supply, except for some gender-specific effects favoring boys.

In Norway, Black, Devereux, and Salvanes (2011) find that young pupils starting school at age 7 rather than 6 have both better cognitive and non-cognitive outcomes at age 18. Andresen (2019, 2018) investigates the effect of the universal early childcare program expansion with large increases of attendance around age 2 on school test scores in reading, math, and English at age 10 for all students in the country. Using an innovative multivalued marginal treatment estimation methodology (see further remarks in section 5), Andresen finds higher enrollment among children from advantaged households. Despite the relatively lengthy gap in time between treatment at ages 1-2 and the measurement of skills at age 10, there are some positive effects of early childcare on school performances. Moreover, there is positive selection associated with larger gains for children with observables which make them more likely to be selected into treatment. Because of the positive selection, further expansion is expected to provide smaller treatment effects than the initial reform, precisely because the children who have the most to gain are already treated. The treatment effect estimates vary across observable covariates and unobserved resistance to treatment, showing substantial heterogeneity in the response to early care, with compliers quite different from the average child.

In Canada, the province of Québec progressively developed, from the end of 1997 over 4 years, for all children 0-4 years not eligible for public kindergarten, a subsidized universal very low-fee (5\$ per day-space) childcare program.<sup>24</sup> This quasi-experimental policy generated several academic studies based on two longitudinal data sets (a very large sample from 1994 to 2008, over 8 cycles with extensive information on child care) representative of all Canadian families with children and youth. Studies on the first decade of the program (Baker, Gruber, and Milligan, 2008, 2019; Lefebvre and Merrigan, 2008, 2009; Haeck, Lebihan, Merrigan, 2018; Kottelenberg and Lehrer, 2013, 2017; Haeck, Lefebvre, and Merrigan, 2015) showed that the province of Québec saw large increases in maternal labor supply and in the enrollment of children in childcare relative to the rest of Canada, where child care policy and services remained largely unchanged. But they all report negative impacts of the policy on children (cognitive, behavioral and health outcomes), as well as parental behavior and well-being in the short and medium run. In the longer run (more so when children are aged 6-9 years and in school), these effects become non-significant, or they fade-out later if significant (Haeck, Lebihan, and Merrigan, 2018; Lebihan et al., 2022). Haeck, Lefebvre and Merrigan (2014), show that the large increases in mothers' labor force participation and earnings impacted the expenditures structure within the household by raising budget shares on expenditures related to children, family goods and services with a collective aspect.

The Canadian province of Ontario, the only one offering a voluntary public school half day pre-K before 2010, introduced a new universal school-year play-based full-day kindergarten program for the 4 and 5 year-olds.<sup>25</sup> Pelletier and Corter (2019) exploited the phasing-in of this program over five years, as a natural experiment in which children from full-day kindergarten could be compared with those from half-day kindergarten in matched neighborhoods. Results showed lasting benefits of full-day kindergarten on children's self-regulation, reading, writing, and number knowledge to the end of Grade 2, including some benefits for vocabulary. Full-day kindergarten children were significantly more likely to meet provincial expectations for reading in Grade 3.

---

<sup>24</sup> At the same time, public school kindergarten for 5-year-old became full-time. New childcare spaces were only added from year 2000. The program was financed by the abolition of many family allowances. Fees were raised over the years in relation with persistent excess demand.

<sup>25</sup> A majority of provinces have non-mandatory full-day kindergarten public school program for the five years old born before December 31 (September 30 in some provinces).

In *Australia*, the impact of universal pre-K on family behavior and child outcomes was estimated for 4-year-olds by examining an experiment where the Australian state of Queensland stopped its public pre-K program in 2007 (Chora et al., 2016). Using a difference-in-differences strategy, they found that five months of access to universal pre-K led to an increase of 0.23 of a standard deviation in general school readiness. Cognitive benefits were displayed across socioeconomic status. The evidence suggests that the positive effects of universal pre-K provision on children's development were driven by the use of higher-quality formal ECE.

### Cross-country studies

There are a few studies using cross-country international data sets on skills achievement, where preschool information and parental education variables are found to play a role in explaining the educational attainment of children.

Comparing the *United States and Denmark* with longitudinal data on test scores, Esping-Andersen et al. (2012) find that enrollment in high-quality formal care at age 3 is associated with higher cognitive scores at age 11, with stronger effects for the very low-income children and for children at the bottom of the test score distribution. But, Datta Gupta and Simonsen (2010), from a pseudo-experiment generating variation in the take-up of preschool across municipalities at age 3, find no significant impact on non-cognitive outcomes at age seven, no matter neither the gender of the child nor the mother's level of education.

Costa and Araújo (2018), from the 2011 PIRLS test scores for three European countries (*Denmark, Sweden, and France*), investigate how achievement relates to student and school with a multilevel analysis including two levels, student/home and schools. Results show that students' home background (home literacy practices, education resources, and reading behavior, all related to parental attributes), and school environment are associated with higher reading scores in all three countries. Hoglebe and Strietholt (2016), with PIRLS 2011 data from *nine countries* and propensity score matching methods, estimate the effect of preschool non-participation on reading literacy in Grade 4. They conclude that even if children who did not attend preschool come mainly from disadvantaged backgrounds in all countries, with the exception of two, reading achievement at the end of primary school is not statistically significantly lower than the performance of matched children from similar backgrounds who attended preschool.

Some studies examine other topics which are loosely related to ECE. For example, Schütz (2009) finds that attendance in pre-primary institutions is positively associated with PISA test scores at age 15 in countries with higher per-pupil spending in pre-primary education. Li and Konstantopoulos (2016), using data from *fourteen European countries* provided by the 2011 sample of the TIMSS survey, present evidence that the effect of class size is generally non-significant for math and sciences different achievement levels. This suggests that in many European countries, class-size reduction as an education policy may not have an impact on student achievement and does not close the achievement gap.

In *Summary*, van Huizen and Plantenga (2018), using meta-regression techniques, analyze the effects of ECE on child development using 253 estimates from 30 studies (from the U.S. and nine other countries) conducted between 2005 and 2015, present overall conclusions on this topic. Children's outcomes – from non-cognitive development measured during early childhood to educational outcomes during adolescence and labor market performance during adulthood – range with effects classified as significantly negative, statistically insignificant or significantly positive. They conclude that the evidence on universal ECE is mixed: neither the age at enrollment nor the program intensity can explain whether the impact is positive or negative. However, quality aspects matters a lot, and some evidence indicate that more intensive programs may produce more favorable outcomes. There is no evidence of effects fading out and the effects of ECE appear to be more favorable in the long run (during adolescence/adulthood). The results show that the gains of ECE are concentrated within the group of disadvantaged children. Evidence on the effectiveness of universal childcare programs targeted towards all children, on the other hand, is also mixed, with effects ranging from negative to positive. One important reason why targeted childcare programs yield larger returns than large scale universal programs may be treatment effect heterogeneity, that is, targeted children from disadvantaged backgrounds may benefit more from the treatment.

**Appendix 2: OLS Estimation of reading z-scores gradients by (1) parental years of education, (2) parental socioeconomic index of occupations, (3) home learning resources scale, (4) preschool duration in years, and jurisdiction, PIRLS 2016**

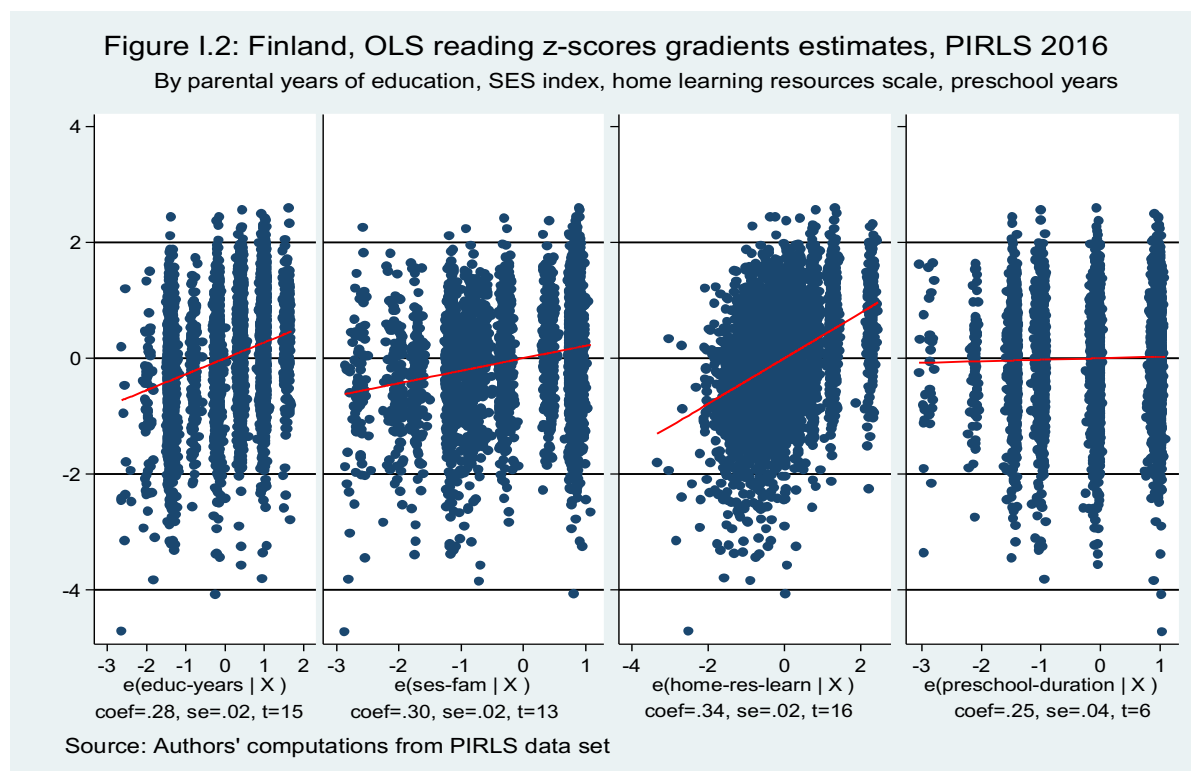
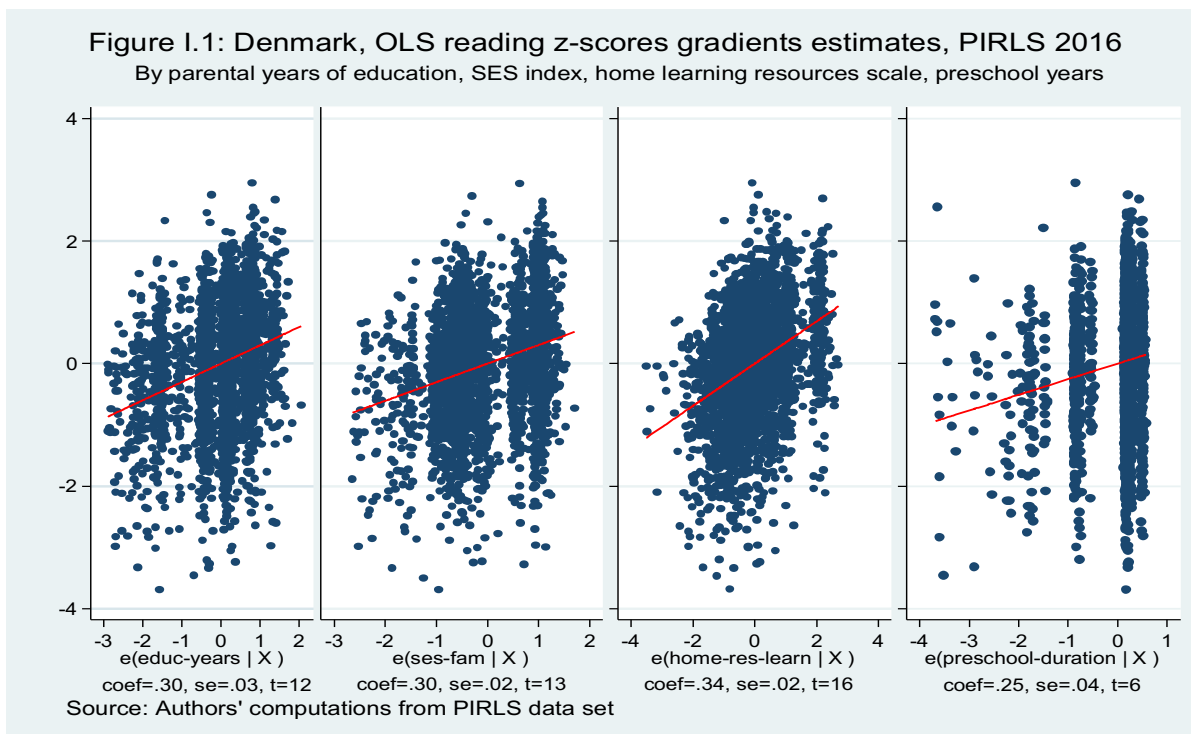
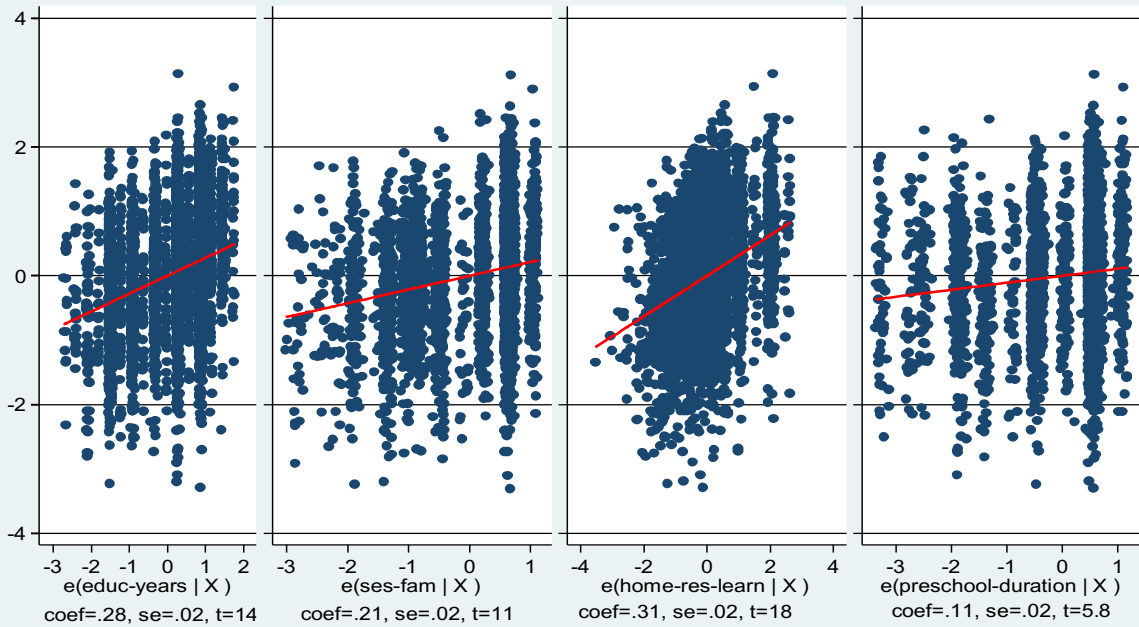
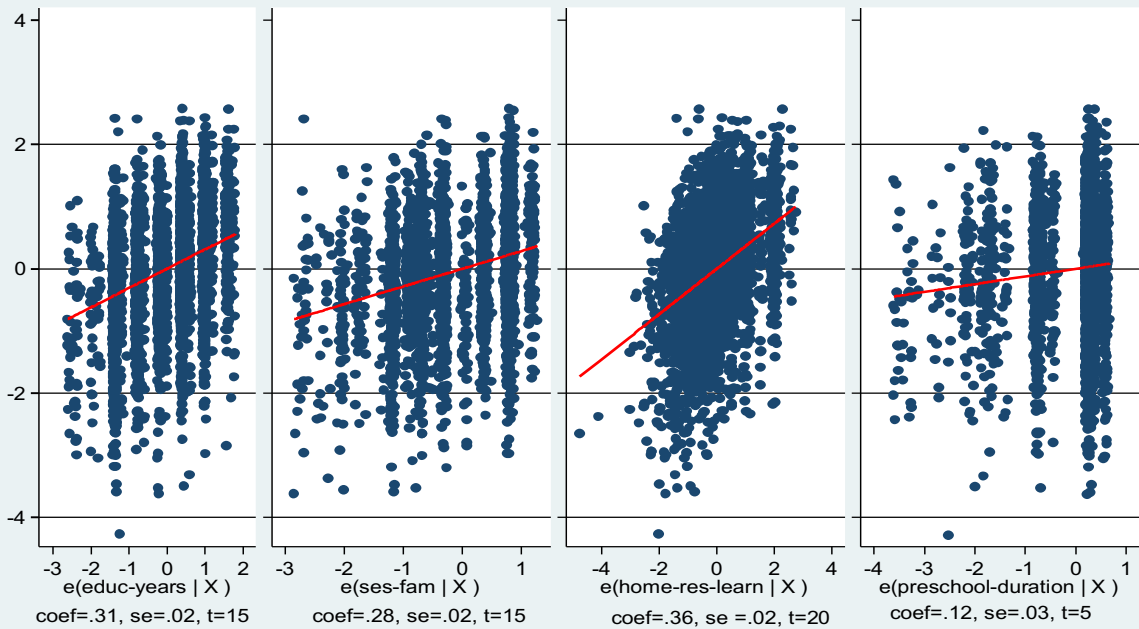


Figure I.3: Norway Grade 5 OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



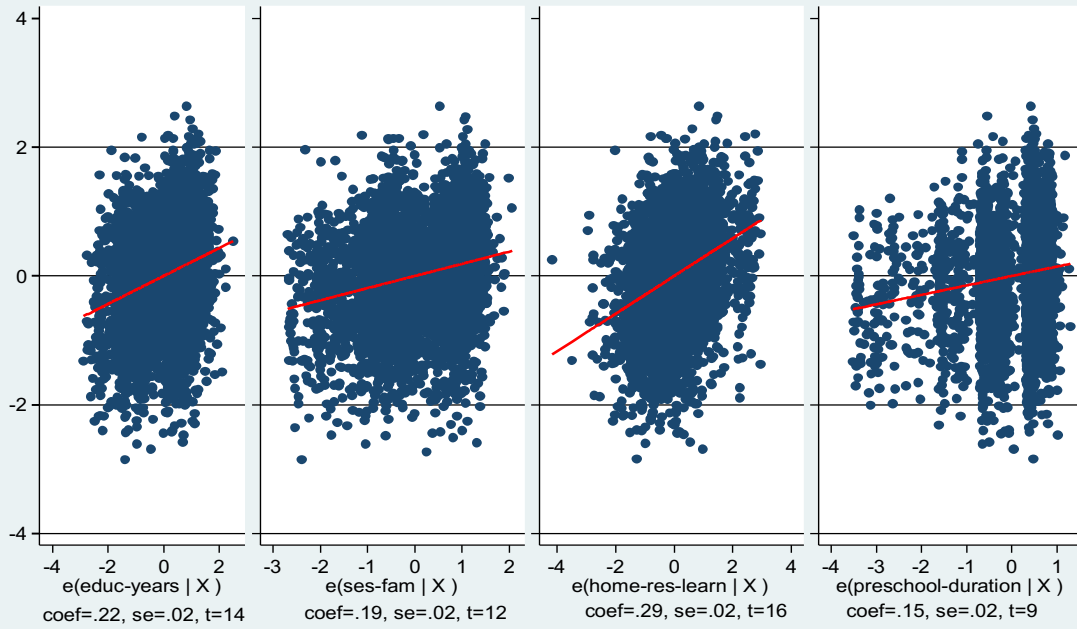
Source: Authors' computations from PIRLS data set

Figure I.4: Sweden OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



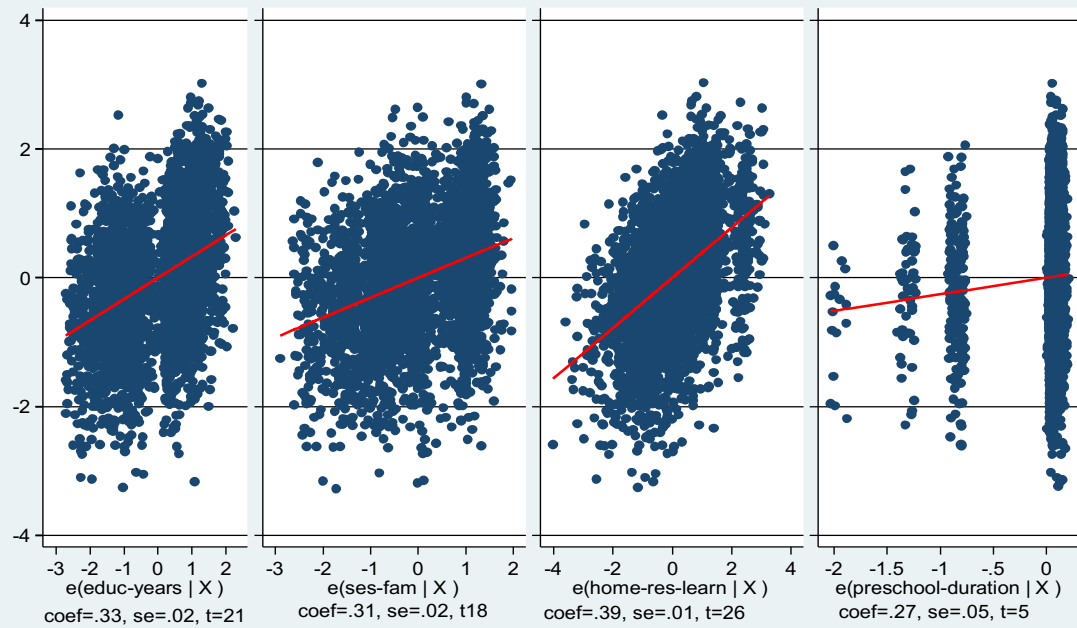
Source: Authors' computations from PIRLS data set

Figure I.5: Belgium Flemish OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



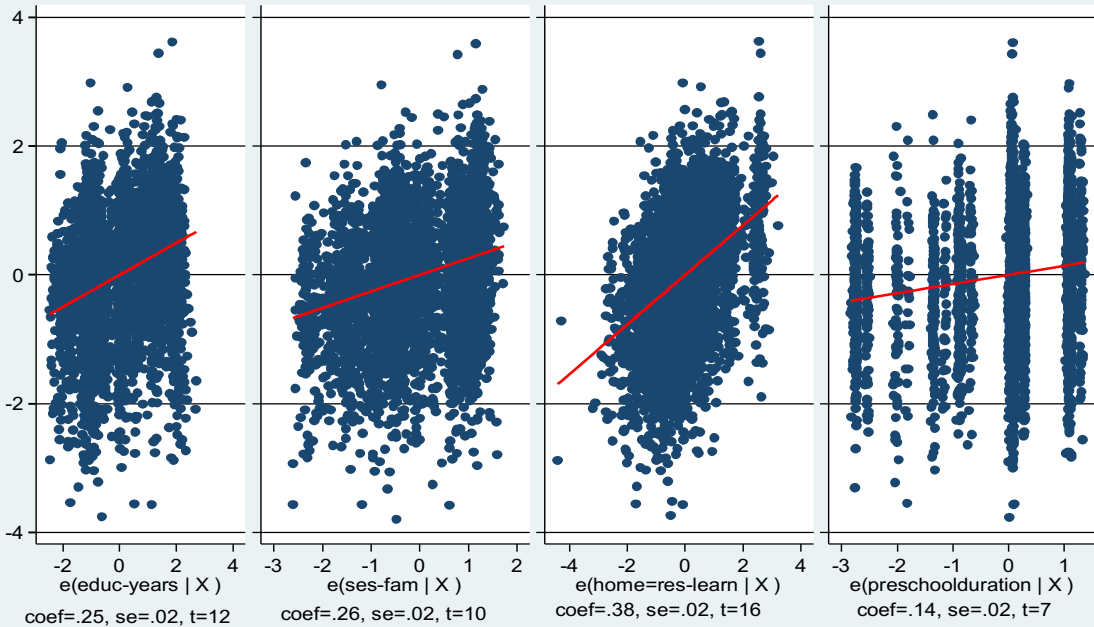
Source: Authors' computations from PIRLS data set

Figure I.6: Belgium French, OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



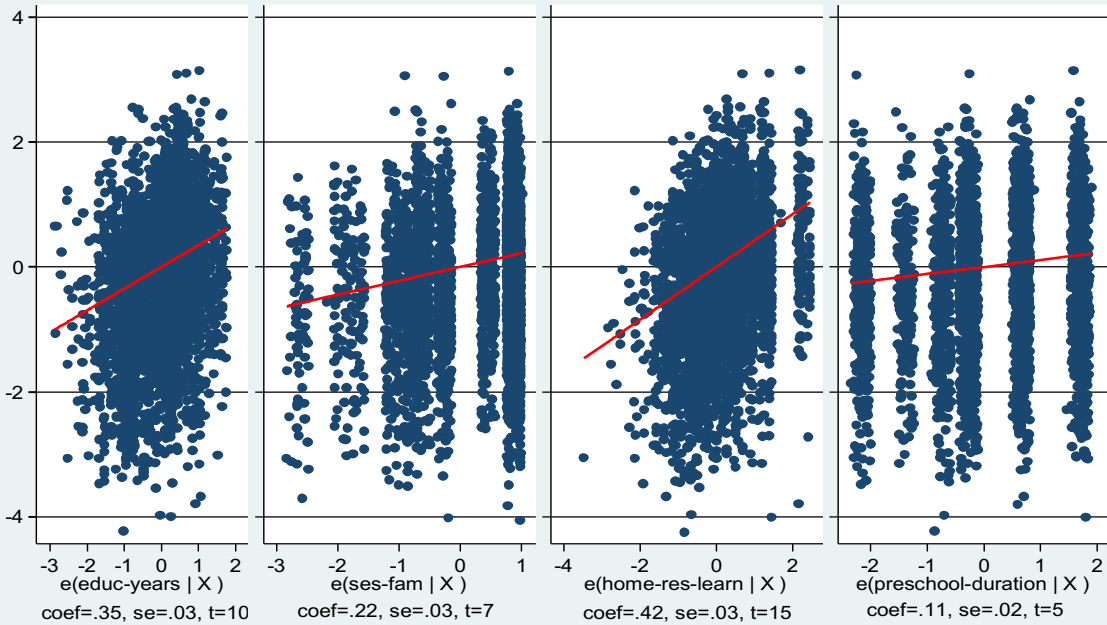
Source: Authors' computations from PIRLS data set

Figure I.7: France, OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



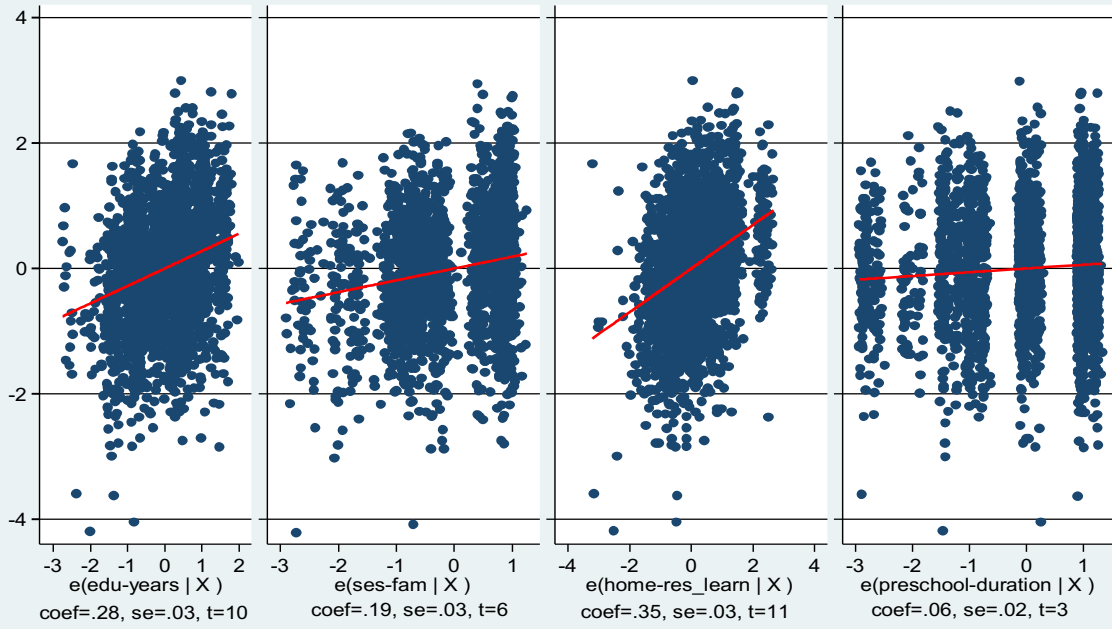
Source: Authors' computations from PIRLS data set

Figure I.8: Ontario, OLS reading z-scores gradients estimates, PIRLS 2016  
 Parental years of education, SES index, home learning resources scale, preschool years



Source: Authors' computations from PIRLS data set

Figure I.9: Québec, OLS reading z-scores gradients estimates, PIRLS 2016  
 By parental years of education, SES index, home learning resources scale, preschool years



Source: Authors' computations from PIRLS data set



Appendix 3: OLS Estimation of math z-scores gradients by (1) parental years of education, (2) parental socioeconomic index of occupations, (3) home learning resources scale, (4) preschool ECE duration in years, and jurisdiction, TIMSS 2019

Figure II.1: Denmark, OLS math z-scores gradients estimates, TIMSS 2019  
By parental years of education, SES index, home learning resources scale, preschool years

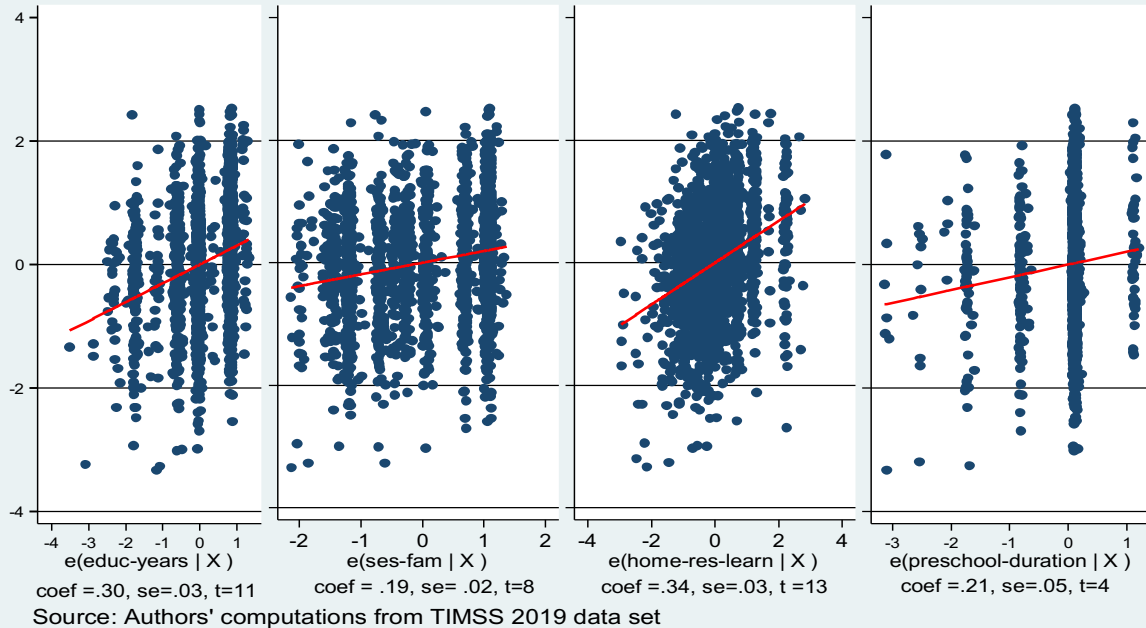


Figure II.2: Finland, OLS math z-scores gradients estimates, TIMSS 2019  
By parental years of education, SES index, home learning resources scale, preschool years

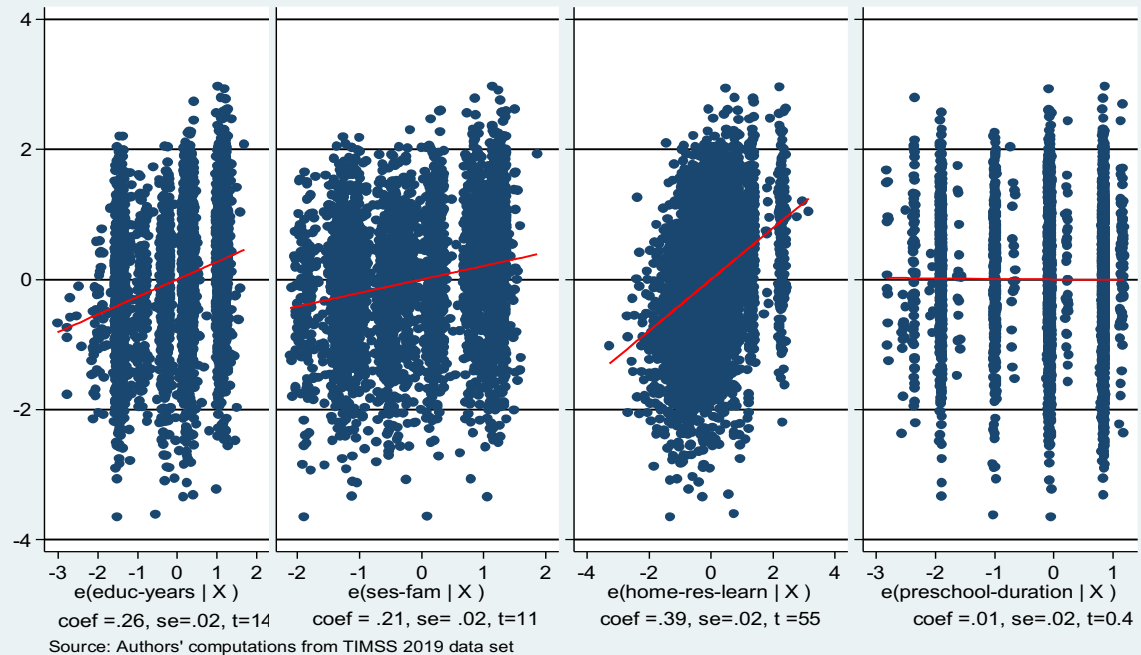


Figure II.3: Norway Grade 5, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index (not available), home learning resources scale, preschool years

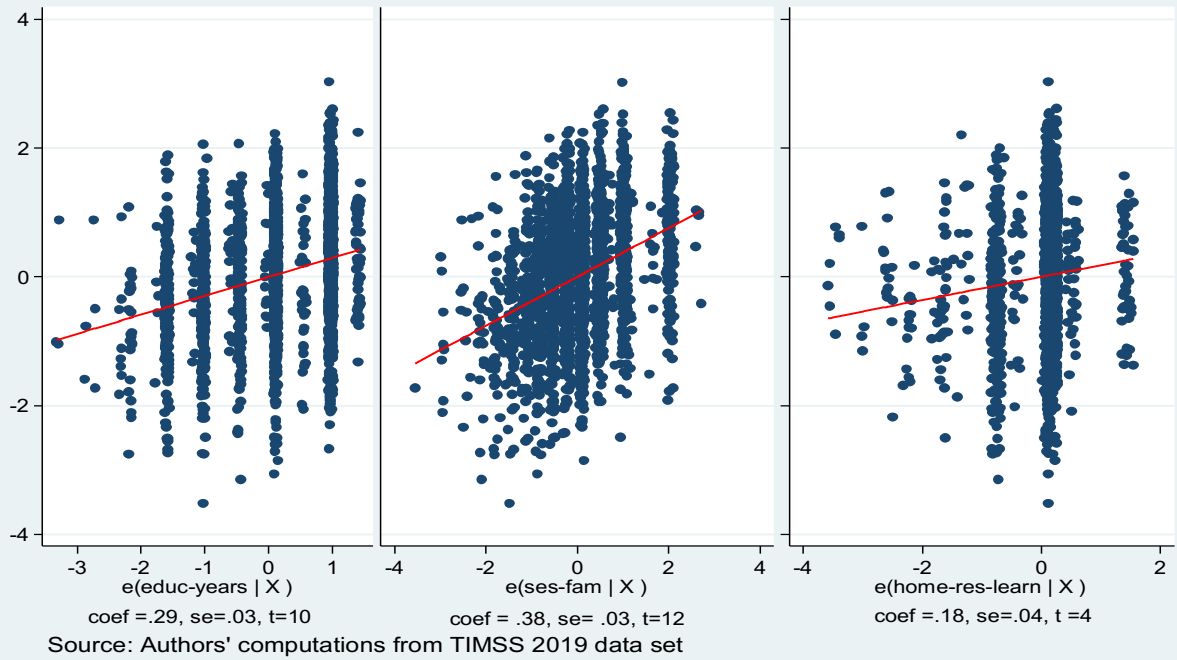


Figure II.4: Sweden, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index, home learning resources scale, preschool years

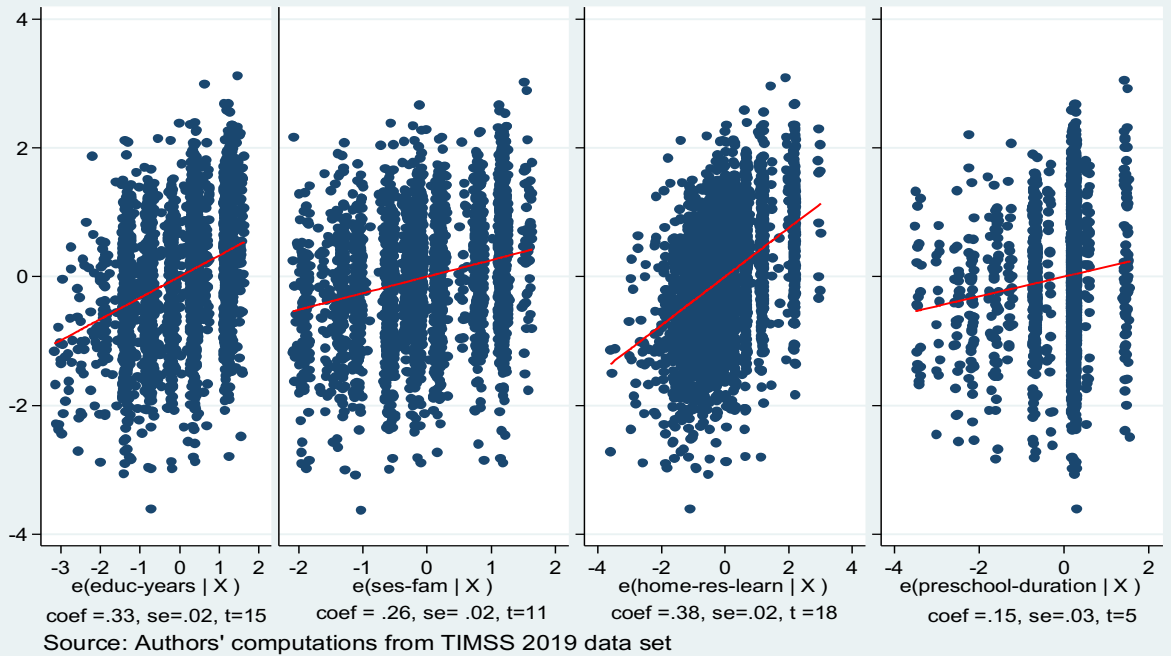


Figure II.5: Belgium Flemish, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index, home learning resources scale, preschool years

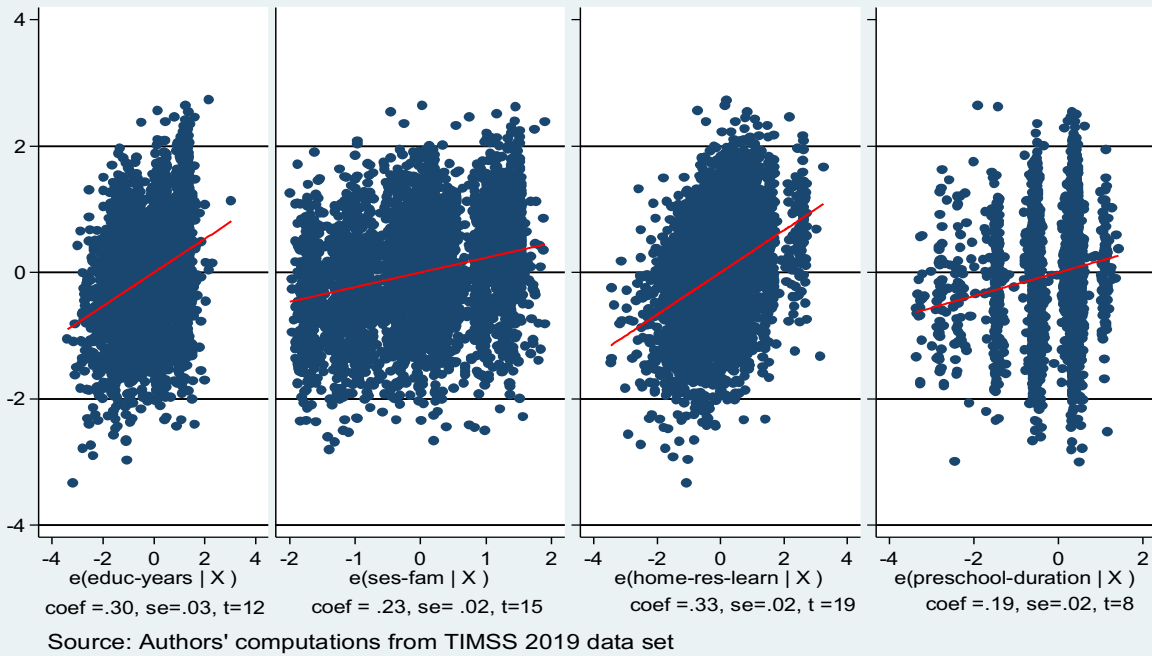


Figure II.6: France, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index, home learning resources scale, preschool years

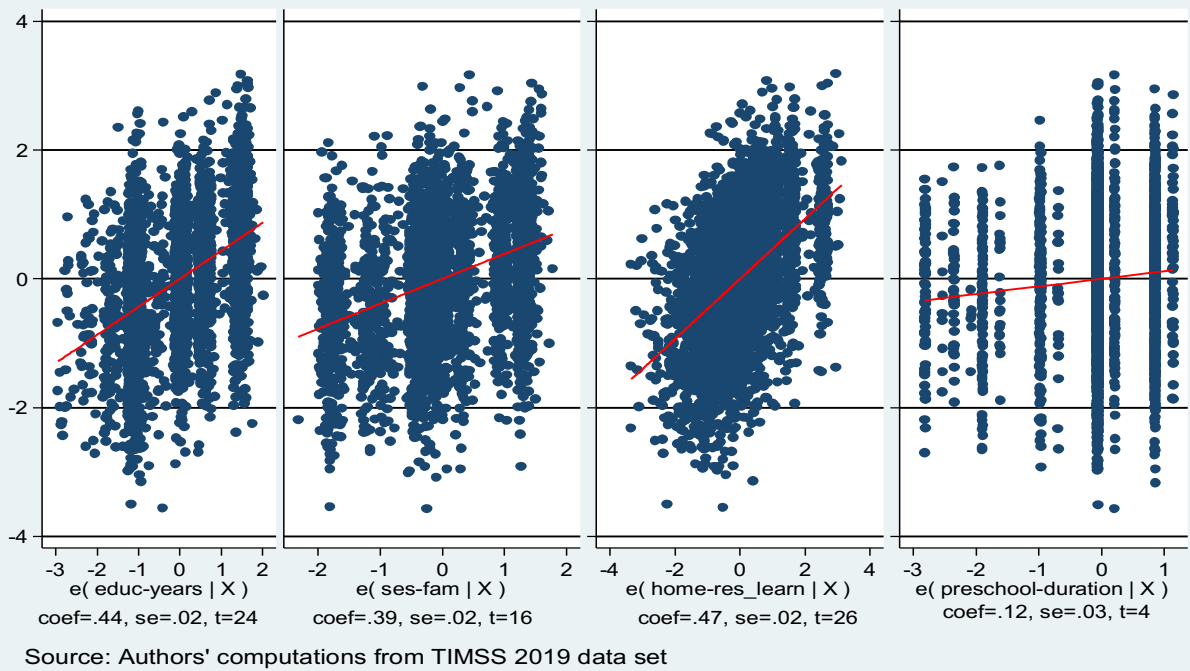


Figure II.7: Ontario, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index, home learning resources scale , preschool years

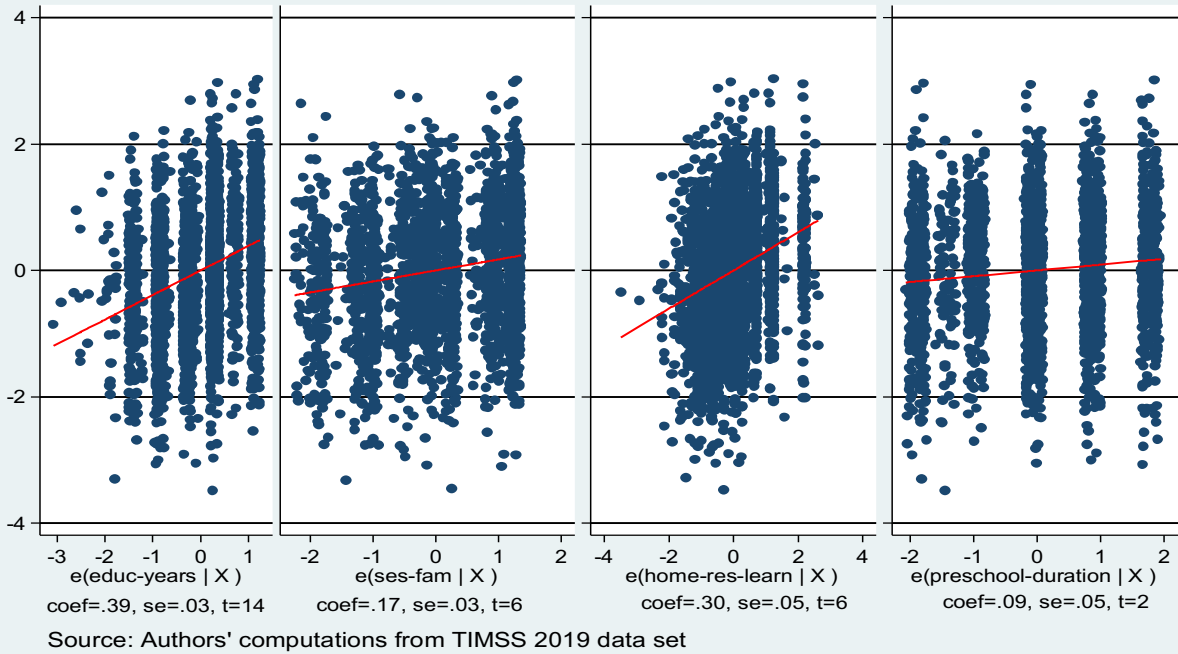


Figure II.8: Québec, OLS math z-scores gradients estimates, TIMSS 2019  
 By parental years of education, SES index, home learning resources scale , preschool years

