I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16956

Public and Parental Investments, and Children's Skill Formation

Miriam Gensowski
Rasmus Landersø
Philip Dale
Anders Højen
Laura Justice
Dorthe Bleses

APRIL 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16956

Public and Parental Investments, and Children's Skill Formation

Miriam Gensowski
Rockwool Foundation Research Unit and IZA

Rasmus Landersø
Rockwool Foundation Research Unit and IZA

Philip Dale
University of New Mexico

Anders Højen
Aarhus University

Laura Justice
Ohio State University

Dorthe Bleses
Aarhus University

APRIL 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Public and Parental Investments, and Children's Skill Formation*

This paper studies the interaction between parental and public inputs in children's skill formation. We perform a longer-run follow-up study of a randomized controlled trial that increased preschool quality and initially improved skills significantly for children of all backgrounds. There is, however, complete fade-out for children with highly educated parents. Given positive long-run effects for children with low-educated parents, the treatment reduces child skill gaps across parents' education by 46%. We show that the heterogeneous treatment effects are a result of differences in parents' responses in terms of investments, reacting to school quality later in childhood. There is also evidence of cross-productivity between reading and math skills and socio-emotional development.

JEL Classification: I24, I28, I21, J24

Keywords: skill formation, parental time investments, public investments, school quality

Corresponding author:
Rasmus Landersø
ROCKWOOL Foundation Research Unit
Copenhagen
Denmark
E-mail: rl@rff.dk

* We are indebted to generous feedback from many, especially Pietro Biroli, Steven Durlauf, Mette Ejrnæs, Sadegh Eshaghnia, James Heckman, Lena Janys, Kristian Bernt Karlson, Claus Thustrup Kreiner, Søren Leth-Petersen, Elena Mattana, Mikkel Mertz, Nicholas Papageorge, Pia Pinger, Almudena Sevilla, Miriam Wüst, and participants at the CEBI/VIVE workshop on "The importance of early-life circumstances: Shocks, parents and policies," Copenhagen Education Network, the University of Copenhagen, the EALE/SOLE/AASLE World Conference, the NBER Summer Institute, the UCL/CReAM Topics in Labour Economics Workshop, and the ASSA 2022 meeting. Miriam Gensowski acknowledges the foundation "A good start in life" for support. The original RCT study of LEAP was supported by a grant from the Danish National Board of Social Services (PI Dorthe Bleses), see https://doi.org/10.1257/rct.6412.

1 Introduction

A large part of inequality in life has its roots in childhood. Therefore, inequality in childhood environments is a major challenge (e.g., Almond and Currie, 2011; Heckman, 2006) that early childhood policies seek to address. Preschools are seen as a main policy tool since there is near-universal enrollment rates by age 4 and 5 in OECD countries (OECD, 2023).

Large and active literatures have established that enrollment in preschool potentially levels the playing field, and that preschool quality matters in this regard.[1] It has also been studied how home and out-of-home environments shape children's skills and inequality across social backgrounds (e.g., Cascio, 2021; Chetty et al., 2014; Cunha et al., 2010; Heckman et al., 2013). However, less is known about how continued effects of improved learning environments in preschool depend on *interactions* between parents' behavior and the school environment that children are exposed to later in childhood. Large gaps exist across socio-economic groups in both parents' investments and children's school quality. Therefore, understanding the interaction between the quality of early childhood environments, parents' behavior, and school quality is crucial for designing policies that effectively reduce inequality of opportunity.

This paper shows how an improvement in the quality of regular universal preschool reduced inequalities in the long-run. We examine potential mechanisms for the heterogeneous effects underlying this equalization, and find that differential parental investments and the quality of the schools that children subsequently attend are likely mediators.

We study a randomized controlled trial (RCT) in Denmark that taught preschool teachers effective practices to support language and pre-literacy skills in children, incorporated in regular classroom activities.[2] The intervention was class-based and not targeted towards specific children. It lasted 20 weeks, but educators were encouraged to continue to use the techniques. Right after the intervention, treated children's language and pre-literacy test scores were increased by approximately 0.3 of a standard deviation, in *all* children irrespective of parents' background. This has been shown previously by Bleses et al. (2018), which we replicate briefly. The homogenous short-

[1] For example, through long-run evaluations of Head Start (Bailey et al., 2020; Carneiro and Ginja, 2014), and studies of the introduction of universal childcare (e.g., Baker et al., 2008; Cascio and Schanzenbach, 2013; Cornelissen et al., 2018; Datta Gupta and Simonsen, 2010a, 2012; Havnes and Mogstad, 2011, 2015). See Blanden et al. (2023) for a comprehensive review.

[2] The RCT was developed by Bleses et al. (2018), who also evaluate its short-run effectiveness.

run effects likely reflect that the intervention focused strongly on scaffolding strategies to help children at their individual base levels of skills. The intervention thereby provides an ideal setting to examine how effects in the longer run depend on later differences in parents' investments and children's schools, potential cross-productivity to other skills, and inequality of skills across family background.

To study long-run effects and potential heterogeneities, we combine information from the original RCT with surveys on parental investments and detailed register data on test scores in grade 2 of elementary school across several skill dimensions, as well as the quality of the children's schools, up to six years later. School quality is approximated by peers' test scores and the characteristics of teachers at the school.

While the intervention resulted in large increases in language and pre-literacy test scores in the short run, we find that by grade 2 (four-five years after the intervention), they fade completely for children with college educated parents. In contrast, treatment effects remain large for children with low educated parents. As in other settings, the intervention takes place against a backdrop of large gaps across parental background in both child skills and well-being, parents' investments in language skills, and the types of schools children attend. The treatment reduced the test-score gap by parental education by 46% — from a baseline gap of almost half a standard deviation.

We also find similar effects on children's math tests and well-being in school. These results suggest cross-productivity across different skill domains.

We present a simple technology of skill formation to highlight the different potential mechanisms that may account for the heterogeneous treatment effects: differences in self-productivity of skills, differences in the effectiveness of parents' investments, and differences in parents' responses to the initial skill improvement from intervention. The model relies on the idea of multiple investments (by parents and schools) — such as the multiple investments studied in Agostinelli et al. (2024). Parents are only assumed to observe their children's skills that were exogenously increased, and re-optimize their investments based on this shock. We find that the third potential mechanism — differences in parents' responses — in combination with the quality of the schools that children attend after preschool, is the likely cause of the heterogenous effects.

Parents increase their investments in response to the intervention only if they are not highly educated *and* if their children attend low quality schools (approximated by their peers' test scores and the characteristics of teachers at the school). This group of parents invests more in both children's reading or language and non-cognitive skills,[3] thereby complementing the initial increase in children's skills from the intervention. The long-run effects on reading ability of children with low educated parents stems almost entirely from the group who attend low-quality schools. In contrast, college educated parents with children in high quality schools *reduce* their investments in their children's non-cognitive skills,[4] and the fade out in treatment effects is also particularly pronounced for this group of children. Overall, the combination of heterogenous treatment effects across both parents' education levels and school quality substantially reduces inequality in child skills in both dimensions.

Our paper contributes to the large literature documenting inequality in children's outcomes across parental background and school or neighborhood characteristics (see Blanden et al. (2023) and Mogstad and Torsvik (2023) for recent reviews). The persistent effects of improved preschool quality (in terms of language learning environment) for children with low educated parents in low quality schools, which are mediated by parents' investment behavior, underscore the importance of complementing early childhood investments with a strong focus on children's home and out-of-home environments in the subsequent years.[5]

Therefore, our paper also contributes to the studies of how parents respond to changes in the quality of educational institutions (Bonesrønning, 2004; Chang et al., 2022; Das et al., 2013; Fredriksson et al., 2016; Gelber and Isen, 2013; Greaves et al., 2023; Houtenville and Conway, 2008; Pop-Eleches and Urquiola, 2013) by studying the differences in parents' investment responses across the quality of the school environment children are exposed to after the treatment. We thereby also add to the literature studying how children's skill formation is shaped by decisions at other margins such as labor supply decisions (Attanasio et al., 2020; Bernal and Keane, 2010; Del Boca et al., 2014; Gormley et al., 2008), school inputs and quality (Ansari and Pianta, 2018; Fort et al., 2020; Nicoletti and Rabe, 2014), children's own investment decisions (Del Boca et al., 2019), the trade off between the quantity and quality of maternal and non-maternal care against mother's labor supply

[3]Throughout the text, we follow the convention in economics of using the term *non-cognitive* skills for self-control, the ability to focus and concentrate, and appropriate socio-emotional development. We appreciate the difficulty that other disciplines would call precisely these skills cognitive skills.

[4]This does not imply that low educated parents invest more than high educated parents. The initial gap in investments across family background is larger than the difference in parents' response to the intervention. The coefficient on reading skills is negative but not statistically significant.

[5]This relates to the study of neighborhood sorting and school choice (see, f.ex. Bayer et al., 2007; Bénabou, 1996).

(Chaparro et al., 2020).

The paper proceeds as follows: Section 2 describes the empirical setting and the intervention. Section 3 describes the data, establishes that treatment and control groups are balanced, and provides evidence of the pre-existing gap in skills across parental background. Section 4 presents the effects of the intervention on reading test scores and other skill dimensions, and Section 5 studies the potential underlying mechanisms. Section 6 concludes.

2 The Empirical Setting

2.1 Preschool in Denmark

Childcare in Denmark is heavily subsidized, and municipalities (local-area governments) are obliged to provide childcare slots for all children (Datta Gupta and Simonsen, 2010b). The typical child starts in nursery at about 12 months of age, and in preschool when they turn 3. Preschool enrollment of all 3- to 5-year-old children is near-universal at 97%. Preschool classrooms comprise on average 20 children with an adult-child ratio of around 1:7 (Slot et al., 2018), with locally determined maxima (European Commission/EACEA/Eurydice, 2019). Most preschool staff (60%) are trained pedagogues with a 3.5 year college degree, and the remainder either have completed short courses in pedagogy or are unskilled.[6] While Danish childcare centers are characterized by high expenditures compared to other countries (Esping-Andersen et al., 2012), substantial sorting by parental income and wealth is already apparent when the child is in nursery (Landersø and Heckman, 2016).

Danish preschools operate from broad "learning schedules," rather than a formal curriculum. The learning schedules are focused on comprehensive personal development, social relations, motor skills, nature and outdoor life, culture, values, relationships, and communication and language (Danish Ministry for Children and Social Affairs, 2018; Slot et al., 2018). There is no systematic and explicit focus on supporting language and pre-literacy skills, reflecting that the main philosophy in preschools is social-pedagogic and not academic (Dahl et al., 2015).[7] Slot (2018) finds that children

[6] Pedagogues' education focuses on children's socio-emotional development. However, to keep a common terminology to other studies in preschool settings, we label preschool staff *teachers*.

[7] In Denmark, nurseries and preschools are placed under the Ministry of Social Affairs (not the Ministry of Education), as childcare was originally introduced as a social policy for low-income and working-class families, and not as an preparation for formal school. There is no distinction between the learning principles in nurseries and preschools with play as a central element (whereas they differ strongly from primary schools).

in Danish preschools generally have higher quality interactions with their peers than preschoolers in countries such as Germany and the U.S, while their interactions with preschool *teachers* are of lower quality. This provides a backdrop in Denmark where there is room for improvement in terms of concerted language-development activities led by preschool teachers.

2.2 The Language and Literacy Intervention

Within this setting, a large-scale language and literacy experiment **LEAP** (Language Education Activities for Preschoolers; *Fart på sproget* in Danish) took place in 2013/2014 with the aim of promoting children's language and pre-literacy skills. There were several non-overlapping treatment arms. See the description in Bleses et al. (2018) for further details on the variations of the intervention in LEAP.

This study focuses on the treatment-arm **LEAP-OPEN**, which varies the quality of teacher's language practices with the children – not other components of the everyday in preschools – and thereby facilitates a test of the paper's central research question: how does an improved preschool environment affect children's skills, parents' investment behavior, and inequality?

The intervention provided teachers with teacher language practices to better incorporate support of language and (pre-)literacy skills with children aged 3–5 in their existing activities. This included, for example, both play-based language and literacy activities (e.g., memory games, rhymes, songs, storytelling using pictures). The teacher language practices were focused on providing rich language and conceptual talk as well as conversational strategies to facilitate dialogue with many conversational turns (e.g., asking open-ended questions rather than closed questions). The material also focused on the use of scaffolding strategies (implemented via a so-called Learner's Ladder) to allow children to receive relevant stimulation of language skills irrespective of their base level (i.e., there was no particular focus on improving skills for specific types of children).

The intervention included preschools from eight municipalities located in different parts of Denmark,[8] constituting more than 1% of all preschoolers in Denmark during that year. The intervention was cluster-randomized at the center level and stratified within municipalities.[9] There

[8]All of the 98 municipalities in Denmark were invited to participate in the project. Out of the 20 that accepted the invitation, the municipalities of Aabenraa, Faxe, Gentofte, Halsnaes, Copenhagen, Lejre, Rudersdal, and Skive were selected to provide a representative sample of the population in Denmark.

[9]Therefore, we cluster at the preschool (i.e., the individual institutions) level in estimations in Sections 4 and 5.

is no overlap between treatment arms in the preschools.

The intervention began with training preschool teachers in the treatment group for two days, which included instruction on the various explicitly stated learning objectives and how they could implement them in instructional activities.[10] The training was conducted by research assistants. During the 20-week intervention, the curriculum was kept open such that the intervention provided teachers with learning targets, teaching material, and a manual, but ultimately allowed each teacher to retain autonomy on how the lessons should be organized and rolled out. Thus, teachers also had autonomy to vary the focus of specific learning objectives, e.g., concrete materials. Preschool teachers in the control group participated in a one-day workshop on topics relevant to the daily routines in a preschool. Parents were *not* treated. They received an initial introduction letter (in both treatment and control groups). Other than that parents were not informed about the activities or intervention.

The intervention's main components were a 20-week period where teachers incorporated activities of their own choice (e.g. shared book reading, nursery rhymes, play-based activities, creative activities etc), while focusing on sequence and scope (targeting specific learning objectives as outlined in Justice et al., 2015; Justice and McGinty, 2012), and scaffolding activities (the Learner's Ladder).[11] The intervention formally consisted of 40 half-hour lessons of high-quality language and pre-literacy learning environment for children. Each lesson was delivered in small groups with around six children and one teacher, with four groups and two teachers per classroom on average in the sample. However, as the intervention was incorporated into the everyday activities, teachers could utilize the intervention's components throughout the day, and they were encouraged to continue using the components after the formal end of the intervention as well.

This intervention constitutes a quality improvement because it introduced a component to the treatment group—systematic focus on language and pre-literacy skill development in conversations and scaffolding strategies targeted to the individual child's need—that is largely absent in the

[10] Two days are, of course, a short time compared to the overall education preschool staff receive. The crucial element, however, is not the length of the course as such, but the fundamental change towards a focus on skills development, including explicit learning targets.

[11] The objectives are concentrated in four domains: i) Vocabulary objectives, *to understand and use words for the names of unfamiliar objects (nouns) and actions (verbs) and that describe things and actions (adjectives and adverbs)*; ii) Narrative objectives, *to identify and describe the setting and characters of a story*; iii) Print knowledge objectives, *to recognize that print carries meaning and to distinguish print from pictures*; iv) Phonological awareness objectives, *to segment words into syllables and to blend syllables into words.*

control group. Importantly, at the same time the intervention can *not* be considered as crowding out other learning activities or replacing a focus from non-cognitive on cognitive language skills. The everyday in Danish preschools is focused on play and socialization activities, with a long-standing focus on stimulating social-emotional skills (Kragh-Müller, 2017). The intervention was placed within this structure. There is no reason to believe that the language stimulation from the intervention crowded out time spent stimulating socio-emotional development.[12]

The average *short-run* effects of the intervention (i.e., after around 6 months) were positive, as presented in Bleses et al. (2018). On average, children from the treatment group scored around 0.3 standard deviations higher than children in the control group in the post-intervention test. We reproduce these results in row 1, column 3 of Table 4 in Section 4.1.

2.3 Primary School in Denmark

In Denmark, compulsory school starts at grade zero (corresponding to Kindergarten in the U.S.) at age 6, and the vast majority of Danish school-age children attend public schools.[13] All Danish public schools have a common curriculum and there is no tracking (by ability, for example) during primary and lower secondary school. Schools are financed by local municipalities, but regulated via a national expenditure rate per student that is based on progressive redistribution between municipalities. Thus, the link between local area public finances and school expenditure is not as strong as in, for example, the U.S. In fact, as Danish schools receive higher rates for special-needs children, the schools with the largest budgets are those with the most disadvantaged students. The distribution of expenditures across public schools in Denmark is very compressed while the U.S. counterpart has large tails both above and below the average expenditure level (see Fig. A.1a).[14]

Nevertheless, parents and teachers still sort based on the characteristics of the students in a

[12]This was explicitly tested in a similar intervention, described in Slot et al. (2018). There, the standardized observation tool CLASS assessed classrooms' emphasis on social-emotional skills in treatment and control groups. Treated classrooms added language activities but maintained the same emphasis on social-emotional skills. Table 5 also tests this by showing effects on math and measures of socio-emotional skills and well-being; there is no sign of the intervention crowding out other inputs to skill formation.

[13]In grades 0 and 1 in 2017, 83% of children attended public schools, and private school enrollment concentrated mainly in schools that cater to religious minorities such as Muslims or Catholics.

[14]Teacher wages in Denmark are set by collective bargaining and schools cannot attract higher quality teachers by increasing wages. Most teachers earn within ±5% of the median wage (see Fig. A.1b), which corresponds to the roughly 5% variation in bargaining for different regions in Denmark to align purchasing power between rural and urban areas. There is virtually no association between teachers' academic skills (proxied by their own high school grade-point-average) and their hourly wages.

given school (see e.g., Eshaghnia et al., 2023), which we will illustrate in Section 3.4 below.

3 Data

Our data set consist of two components: Intervention data with information on treatment assignment and short run test scores, and register data with background information and longer run test scores. Full details on the data construction are given in Appendix C.

3.1 Intervention Data

The intervention data includes information on treatment status, unique preschool identifiers, and unique individual identifiers allowing us to link all children from each preschool. The individual identifiers also facilitate the link to the register data, which we describe below. Language tests were collected before and after the intervention (child ages 3-6), mostly by the staff in the preschools, and the tests follow the official language screening used in Danish preschools (Bleses et al., 2018; Højen et al., 2022).

We exclude immigrants from our main sample, as the study focuses on inequality between children, and how parental and public investments in their skills mediate the effects of an early *language* intervention.[15] The role of parental language investments is fundamentally different if parents themselves do not speak the language that is taught in the institutions. Most importantly, it is natural to have a conceptual idea of child-led change: Parents observe their child's improved skills and react to those. Parents not speaking Danish might not easily gauge their child's language level. Therefore, any reactions these parents have following the RCT would likely come from very different mechanisms than what would be the case for Danish parents.

In total, 2,515 children from 73 preschools were part of the intervention, of which 2,301 have both valid pre-trial and post-trial language test scores (the missing observations relate to e.g., absence at the day of assessment) and are included in our main sample.

[15] To assuage any doubt, we perform robustness checks with the full sample that includes children with an immigration background. Reassuringly, Table A.21 shows that treatment effects on language test scores remain very similar. The treatment effects are not statistically different between native and non-native children (Table A.22).

3.2 Follow-up Survey

Around 3 years after the intervention, we conducted a follow-up survey of the parents. The invitation to participate in the survey was sent via secure email to both treatment and control groups. Parents who did not complete the survey within 10 days were subsequently contacted via telephone. We obtained a response rate of 58%. The sample size with parental survey responses is 1,338. Attrition or selection into response was not influenced by treatment status (Table A.3).

In the questionnaire, we asked parents about activities relating to support of their children's development in the language/reading domain as well as for socio-emotional skills. We construct a measure for **parental reading investments** and **parental non-cognitive investments** in their children with confirmatory factor analysis. The reading measure is based on items such as "I enjoy reading for my child.", "I am often too busy or too tired to read to my child.", and "How many times last week has your child been read to (or read with) at home?".[16] Example items for the non-cognitive investments are "I do a lot to teach my child to focus, concentrate, and complete a task." or "When I play or read with my child, it is important to finish before we stop or start new things." See Section C.1 for more information about the survey, Section C.4 for the full list of questions, and Tables C.1 and C.2 for the factor loadings.

3.3 Register Data

We link each child and their parents to full population register data using the unique individual identifier of the child's social security number. This link yields a data set on not only the intervention but also the children's long-run outcomes as well as their family's and school's characteristics.

Child skills measured in school: We obtain a follow-up measure of child skills from compulsory national tests in grade 2 (age 8–9, which is 3–5 years after the intervention), that assess children's language comprehension and reading skills (decoding ability and reading comprehension). Answers to all domains are aggregated based on a Rasch model that minimizes the uncertainty of each student's performance assessment. For readability, we refer to the combined score as **reading test**

[16] Since some of the items could arguably more related to time investments whereas others could reflect enjoyment of these activities, we also analyze a split version of this factor. See Table A.8 for these results.

scores (grade 2). The tests are computerized, objective, and fully comparable across schools.[17]

To test cross-productivity of different domains of child skills, we add **mathematics test scores** in grade 3 from the same type of national tests. These cover numbers and algebra, geometry, statistics and probability. We also measure **socio-emotional skills** through the well-being survey in grade 2. For this domain, we extract three factors from 18 items that each have 3 answer categories (see all factor loadings in Table C.2).[18] **General well-being** is at the intersection between school-well being and academic functioning, as items loading on this factor are not only "Are you happy with your school?" and "Do you learn something interesting in school?" but also "Can you concentrate in class?". Items that mainly load on **social skills** are "Do you like the breaks at your school?", "Do you think the other children in your class like you?". Example items for **socio-emotional distress** are "Do you feel alone at school?", "Does your stomach hurt when you're in school?" and "Are you afraid the other children will laugh at you?"

All of the tests and surveys that assess children's long-run skills in our setting are mandatory for public schools to administer. They are not necessarily mandatory for individual children to sit (if they are ill, for example), and will be missing for children attending some private schools.

Other characteristics: The register data also include unique links to parents' individual identifiers allowing us to link each child to their family, home address, and parents' income, employment status, and highest educational attainment. Degrees attained are converted to years of education by the standard length for obtaining a given degree. We proxy parental skills by their education, where we categorize them by their highest recorded years of schooling, classifying parents where none has at least 14 years of schooling as "High school or less" (36% of our sample), and those where at least one has 14 or more years as "College or more." We refer to the two groups as *low educated* and *college educated* parents for simplicity.

The full list of covariates we use as controls are child's age, gender, birth weight, gestational length, APGAR score, and number of siblings, mother's weight at child's birth, and each of parents' age, years of education, employment status, as well as household wage income and dummies for

[17]The tests take place near the end of the school year and are computerized adaptive tests in which questions are determined by the student's performance earlier in the test. The tests are scored electronically without teacher input. Following Sievertsen et al. (2016) and Beuchert and Nandrup (2018), we standardize these three individual scores, take the simple average, and re-standardize them within year.

[18]For older students, starting in grade 4, there is a 40-item survey that contains items that can be used to extract personality traits (see Andersen et al., 2020). This is not possible here.

child's birth year.

School quality: Our main measure of school quality is the average performance of students at a given school on compulsory (externally scored) national tests for grades 2–8 (Danish and math) in years 2010–2016.[19] Abdulkadiroğlu et al. (2018) explicitly find that parents value peer *achievement*, not value added, and other studies also find that parents care about a school's test scores as a marker of quality (see, for example, Burgess et al., 2015, and the work they cite). We define each child's school as the school they attended when they completed their grade 2 reading test, and we rank the quality from 0 to 1, 0 being the school with lowest average test scores and 1 the highest.[20] When we later split the analysis by "School quality low" vs high, we cut at the 40th percentile of the population. In our sample, this corresponds to the bottom quartile vs the top three in "School quality high."

Analysis Sample: Our main analysis sample for the immediate effects of the intervention consists of 2,301 Danish children with test scores and information on parental education. There are 1,151 children in the treatment group (36 institutions) and 1,150 children in the control group (37 institutions), respectively. Longer-run reading tests in grade 2 are available for 1,898 children. Data availability is not influenced by treatment status, as we test in Table A.3 for all longer-run outcomes (grade 2 reading, grade 3 math, the socio-emotional measures and the parent survey).

Table 1 presents an overview of when the main variables of our analysis were measured, and

[19]Our main measure of school quality captures *peer quality*. We also define an alternative measure of school quality, where we predict how well students at each school should perform on the national tests, based on average *teacher* characteristics at this school. For this, we use a unique dataset linking all teachers in Denmark and the schools at which they work at from 2010 to 2016. We use individual teachers' information on age, tenure, year of graduation, high school grade-point-average (GPA), high school GPA in language subjects, GPA from teachers' college, and unemployment spells, and calculate the school averages of these teacher characteristics. We then link each school with the individual standardized language test scores for all students from grades 2 through 8 during the years in question, and regress children's test scores on the average teacher characteristics in the school they attend. The results from this association of child outcomes with school (teacher) characteristics are then used to predict each school's expected average child outcomes. The correlation between the two measures is 0.71, see Table 3, and results based on the two measures are also very similar, see Table A.12.

[20]Thus, we measure school as the realized school choice, which could lead to biased results if parents change sorting behavior following the intervention: While access to a specific school is determined via school catchment areas (set by home-addresses), there are a number of open slots for children from outside the catchment area in most schools, and access via this channel tends to be a function of parents' efforts. Bjerre-Nielsen and Gandil (2020) show that affluent parents often succeed in placing their children in local public schools with other high-SES children, even when this means going out of their own default catchment area. To test this, we also estimate effects using the *default school* based on the modal school choice in the population of students at each preschool that participated in the intervention. While there is some variation between realized and default school, results are very similar irrespective of what measure we use (see Table A.13). We therefore use realized schools to define quality.

what model components they relate to. The pre-trial test scores were collected a few weeks before the intervention, and the post-trial test scores shortly after. Then, in 2016, 3 years after the intervention, the parent survey was collected. In 2017-2020, we measure children's reading test scores (when they are in second grade), as well as their schools' quality.

Table 1: Timeline of outcome measurements

Timing	Age	Event	Data
-3 – -6 months	3–5	Baseline data collection	Language test scores
0 – +5 months	3–6	*Intervention*	
+10 months	4–7	Endline data collection	Language test scores
+3 years	6–8	Parent survey	Parental investments
+3 – 6 years	7–9	In school tests	Reading and math test scores / Socio-emotional skills / School quality

Note: Timing is relative to the intervention.

3.4 Descriptive Statistics and Balancing

We present average characteristics of the full population of children born in the same years as our sample in Table 2, column 1, and the corresponding characteristics for our sample in columns 2 and 3 (treatment and control group, respectively). The table shows that our sample is representative of the average child in Denmark with only a few exceptions: Parents in our sample are marginally older, have 0.2 years more schooling, and the mothers have higher employment rates compared to the average mother in the overall population.

Importantly, columns 2-4 in Table 2 show that children's pre-intervention test scores, own background characteristics, and family background characteristics are balanced across treatment assignment. Of the 19 covariates only one is significantly different at the 10% level when we test for differences in the treatment and control group in column 4. In a joint test where treatment status is regressed on all the covariates for the full sample and separately for children with low and college educated parents, two of the 57 tests are significant at a 5% level and four at a 10% level (Table A.1). The results from the balancing tests correspond to what could be expected from random variation. Table A.2 reports similar balancing test when we consider the sub-sample with grade 2 reading test scores, grade 3 math tests scores, grade 2 well-being survey, and parental responses to the follow-up survey, Table A.3 tests for association between data availability and treatment status, and Table A.4 present the association between data availability and covariates. There is no sign of selective attrition by treatment status, which otherwise could have affected the credibility of our research design.

Table 3 shows the correlations between the main measures used in the paper. While the language test scores taken at different points in time are significantly correlated, they are not near-identical with correlations of 0.52 and 0.29 between the pre-trial test score on the one hand and the post-trial and grade 2 test score on the other. Language test scores are also significantly correlated with math test scores in grade 3 and social skills in grade 2. The table further shows that children's test scores are correlated with both parents' years of schooling, and that this association increases as children age (ρ of about 0.2 in the preschool tests and 0.3 in the grade 2 test).

Children's test scores are also positively associated with the different measures of school quality – both when tests are taken before and after school starts (ρ of about 0.1 and 0.25, respectively). Finally, the table shows that parents' educational attainment is positively correlated with school quality (ρ of approx. 0.3) measured either through average test scores at the school or teacher characteristics. These baseline associations illustrate that both families and teachers sort based on the student body at a given school.[21]

To obtain a grasp of the extent of the baseline inequality that the intervention is up against, we examine children's skills by parental education *in the control group only*. Figures 1a and b present the post-intervention language and grade 2 reading test score distributions by parents' education level. The figures document stark differences between children's language skills across parents' education levels — even at an early age. Moreover, the differences across parental background are present throughout the skill distribution and not just in specific ranges. Similarly, Figure 1c presents the grade 2 reading test score distribution separately for children attending low- and high-quality schools with clear differences in test scores across school quality throughout the distribution.

[21] Landersø and Heckman (2016) and Eshaghnia et al. (2023) illustrate the sorting in Danish schools.

Table 2: Balancing of Estimation Sample by Treatment Status

	(1) General Pop	(2) Control Avg.	(3) Treated Avg.	(4) Diff Treat-Control
Pre-Trial Test		0.075	0.002	0.073
		(1.019)	(0.998)	(0.088)
Child Age at pre-trial test		4.041	4.086	-0.045
		(0.851)	(0.849)	(0.044)
Born in 2007	0.253	0.174	0.190	-0.017
	(0.434)	(0.379)	(0.393)	(0.046)
Born in 2008	0.256	0.330	0.331	-0.001
	(0.436)	(0.470)	(0.471)	(0.024)
Born in 2009	0.246	0.331	0.295	0.036
	(0.430)	(0.471)	(0.456)	(0.022)
Born in 2010	0.246	0.156	0.177	-0.021
	(0.431)	(0.363)	(0.382)	(0.043)
Male	0.514	0.504	0.537	-0.033
	(0.500)	(0.500)	(0.499)	(0.026)
Birth weight (kg)	3.478	3.461	3.517	-0.056**
	(0.604)	(0.530)	(0.484)	(0.025)
Gestation (wks)	39.627	39.703	39.771	-0.067
	(1.945)	(1.597)	(1.565)	(0.080)
Apgar score	9.863	9.876	9.852	0.024
	(0.620)	(0.594)	(0.591)	(0.029)
Number of Siblings	1.382	1.322	1.389	-0.066
	(0.851)	(0.785)	(0.860)	(0.058)
Mother weight (kg)	67.569	67.545	66.515	1.030
	(38.924)	(15.911)	(16.804)	(1.277)
Mother education (yrs)	13.934	14.115	14.185	-0.071
	(2.457)	(2.436)	(2.514)	(0.306)
Mother age	39.153	39.823	39.889	-0.066
	(5.050)	(4.971)	(5.103)	(0.553)
Mother employed	0.788	0.844	0.830	0.015
	(0.409)	(0.363)	(0.376)	(0.025)
Father education (yrs)	13.678	13.861	13.948	-0.087
	(2.438)	(2.422)	(2.422)	(0.296)
Father age	41.532	42.049	42.457	-0.408
	(5.777)	(5.629)	(6.006)	(0.524)
Father employed	0.882	0.898	0.900	-0.002
	(0.322)	(0.302)	(0.300)	(0.019)
Household inc(1,000 USD)	89.241	99.358	100.194	-0.836
	(58.087)	(61.277)	(72.184)	(8.317)
School Quality	0.640	0.634	0.644	-0.010
	(0.247)	(0.240)	(0.254)	(0.054)
School Quality (Teacher Characteristics)	0.509	0.503	0.515	-0.011
	(0.317)	(0.315)	(0.319)	(0.075)
Observations	235,194	1,150	1,151	2,301

Note: The table shows descriptive statistics for all non-immigrant children in Denmark in the same birth cohorts as the paper's sample (column 1), the control group (column 2), and the treatment group (column 3). Average treatment-control differences are shown in column 4. Standard deviations of the variables are shown in parentheses for columns 1-3, standard errors clustered at institution level for column 4. The general population (column 1) consists of all children born in 2007-2010. $^{*}p < 0.10$,** $p < 0.05$,*** $p < 0.01$. Note that the number of observations is 2,036 for the two rows of school quality.

Table 3: Correlation Among Key Variables — Control Group

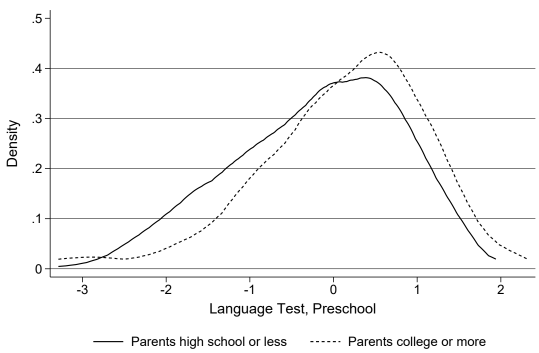
	Pre-Trial Test	Post-Trial Test	Reading (Grade 2)	Math (Grade 3)	General well-being	Social skills	Socio-em. distress
Pre-Trial Test	1.000						
Post-Trial Test	0.521***	1.000					
Language Test (Grade 2)	0.290***	0.277***	1.000				
Math (Grade 3)	0.290***	0.286***	0.510***	1.000			
General well-being	−0.020	−0.035	0.097**	0.061	1.000		
Social skills	0.051	0.029	0.152***	0.133***	0.303***	1.000	
Socio-em.distress	−0.095**	−0.094**	−0.074*	−0.073*	−0.453***	−0.357***	1.000
Reading Inv.	0.107***	0.050	0.062	0.007	0.052	0.070	−0.117**
Non-cog Inv.	−0.050	−0.071*	−0.102**	−0.106**	0.067	−0.072	−0.003
School Quality	0.090***	0.089***	0.236***	0.214***	0.076**	0.144***	−0.077**
School Quality (Teacher Characteristics)	0.133***	0.116***	0.244***	0.208***	0.046	0.129***	−0.076
Mother education (yrs)	0.212***	0.175***	0.315***	0.323***	0.096**	0.096**	−0.134***
Father education (yrs)	0.185***	0.113***	0.287***	0.301***	0.020	0.074*	−0.080**

	Reading Investment	Non-cog. Investment	School Quality	School Quality (teach.)	Mother education	Father education
Reading Inv.	1.000					
Non-cog Inv.	0.151***	1.000				
School Quality	0.036	−0.065	1.000			
School Quality (Teacher Characteristics)	0.026	−0.031	0.713***	1.000		
Mother education (yrs)	0.096**	−0.106***	0.367***	0.297***	1.000	
Father education (yrs)	0.042	−0.079**	0.401***	0.323***	0.497***	1.000

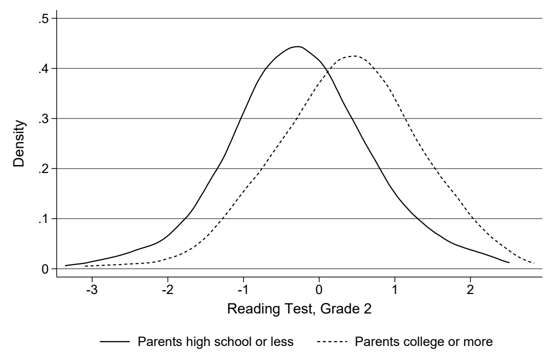
Note: The table shows pairwise correlations (i.e.,not restricting on a balanced sample) for all key outcomes and characteristics considered in the paper. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Figure 1: Baseline Differences: Children's Language Scores and Parents' Investments in Control Group, by Parents' Education

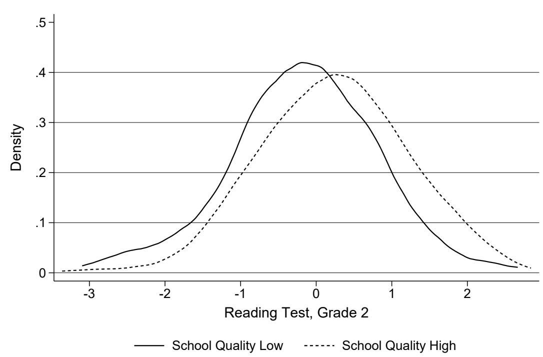
a) Language Test Scores, Post Trial
by Parents' Education



b) Reading Test Scores, Grade 2
by Parents' Education



c) Reading Test Scores, Grade 2
by School Quality



Note: The figure shows test score distributions (for the control group) by parents' education and school quality. Fig. a) shows post-trial test scores and Fig. b) shows grade 2 test scores by parents' education. Fig. c) shows grade 2 test scores by school quality.

4 Results

4.1 Direct Effects on Language Skills

As a first visual illustration of the treatment effects on children's language test scores, Figure 2 shows the distributions of language test scores measured before the intervention, shortly after the intervention, and in grade 2, by treatment status and parental education. Before the intervention, there is balancing: no treatment-control differences in pre-trial language test scores. Shortly after the intervention, the distribution of test scores is shifted rightwards in the treatment group — for children of both highly and less educated parents. By grade 2, any treatment-control differences have vanished for children whose parents have a college degree, while the clear differences remain for children with parents who have no more than a high school education.

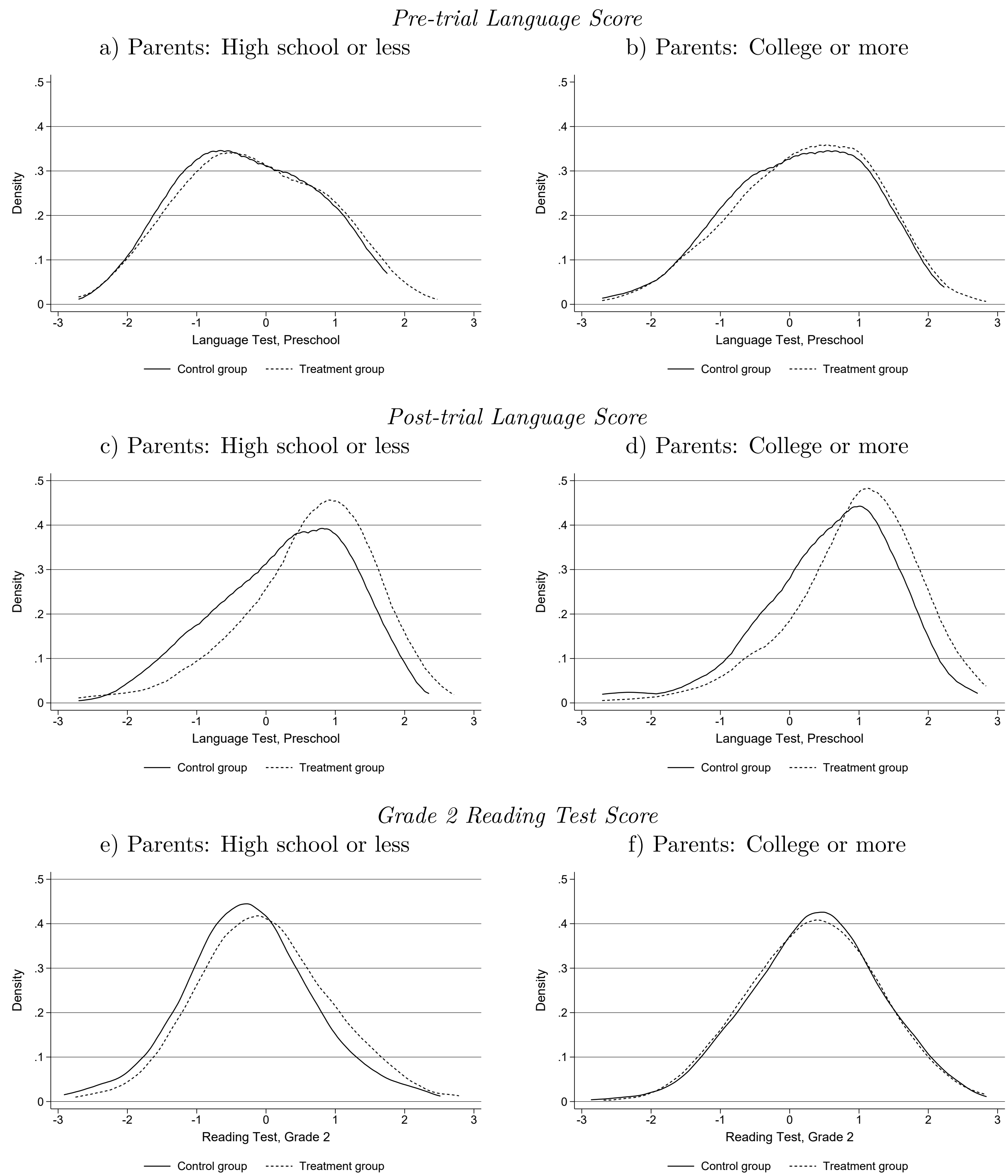
To quantify these differences, we estimate the average effects of the intervention as:

$$y_{it} = \alpha_t + \beta_t T_i + \gamma_t X_i + \varepsilon_{it} \tag{1}$$

where y_{it} are child i's language and reading scores at the different time points t, T_i is the treatment indicator, and X_i are control variables (child's age, gender, birth weight, gestational length, APGAR score, number of siblings, mother's weight at child's birth, and each of parents' age and employment status, as well as household wage income and dummies for child's birth year). We also include parental years of education unless we test for heterogeneous effects by high school/less vs. college/more (where these indicators are included instead). The standard errors are clustered at the preschool level.

The first row in Table 4 shows the estimated treatment effects, β_t based on Eq. (1). Column 1 shows that there is balancing at baseline: no differences in test scores before the intervention by assigned treatment status. However, turning to test scores measured after the intervention, column 3 shows that the intervention is followed by a large treatment effect of 30% of a standard deviation. This effect is around the same size as the average effects of teacher coaching interventions surveyed in the meta analyses by Kraft et al. (2018) and Markussen-Brown et al. (2017). Yet, by grade 2

Figure 2: Test Score Distributions, by Treatment Status and Parental Education



Note: The figure shows pre-trial, post-trial and 2nd grade test score distributions for the control and treatment group by parents' education, in standard deviations from the mean. Parents where neither of them has at least 14 years of schooling are classified as "High school or less" vs "College or more." Note that the samples vary from the two pre-and post-trial tests to the 2nd grade reading test score, see Table 4.

(column 5), treatment-control differences appear to have faded.[22]

Table 4: Baseline Balancing, Short-run, and Longer-run Treatment Effects

	Pre-Trial Test		Post-Trial Test		Reading Test Grade 2	
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	By parent edu	Pooled	By parent edu	Pooled	By parent edu
Treated	0.084		0.309^{***}		0.055	
	(0.076)		(0.058)		(0.076)	
Treated × High school/less		0.079		0.304^{***}		0.222^{**}
		(0.095)		(0.065)		(0.095)
Treated × College/more		0.091		0.311^{***}		−0.039
		(0.089)		(0.069)		(0.076)
College/more		0.324^{***}		0.075		0.483^{***}
		(0.073)		(0.051)		(0.070)
Parental Education	X	-	X	-	X	-
Pre-test	-	-	X	X	X	X
Covariates	X	X	X	X	X	X
Observations	2,301	2,301	2,301	2,301	1,898	1,898

Note: Regression estimates of the treatment–control differences (β_t) in test scores y_{it} from $y_{it} = \alpha_t + \beta_t T_i + \gamma_t X_i + \varepsilon_{it}$, and with an interaction of treatment status T_i with parental educational attainment. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Main analysis sample excluding children with immigration background. Other covariates included are child's age, gender, birth weight, gestational length, APGAR score, number of siblings, mother's weight at child's birth, and each of parents' age and employment status, as well as household wage income and dummies for child's birth year. Parental years of education are included as covariates only in the pooled regressions. See results for these coefficients in Table A.5.

The fact that the sizeable initial treatment effect fades out by grade 2 may cover heterogeneity across parental background. We thus estimate the effects of the intervention by parents' education:

$$y_{it} = \alpha_{0,t} + \alpha_{1,t}\text{college}_i + \beta_{0,t}T_i \cdot \text{HS}_i + \beta_{1,t}T_i \cdot \text{college}_i + \gamma X_i + \epsilon_{it} \tag{2}$$

where college_i and HS_i are dummies indicating whether child i's parents have completed at least a college degree or not (high school or less). The estimates of $\alpha_{0,t}, \alpha_{1,t}$ will capture the average of y_{it} in the control group by parents' education level, and the estimates of $\beta_{0,t}, \beta_{1,t}$ will capture the average treatment-control differences by parents' education level.

Column 4 in Table 4 shows that the intervention was equally effective for both groups of parental

[22]The results are robust to the exclusion of covariates or pre-test scores in the conditioning set, which is expected given the balancing tests presented in Section 3. See Table A.15

education: short run average treatment effects (post-trial test) are similar with estimates of 0.30 and 0.31 of a standard deviation for low educated and college educated parents, respectively. But by grade 2 (column 6), the effects have faded completely for children with college educated parents, while effects remain sizeable at 0.22 of a standard deviation for children with low educated parents. This effect corresponds to 46% of the baseline test score gap across parental education. The difference between low and college educated parents are statistically highly significant.

One potential explanation of the fading effects for children with college educated parents is a *ceiling effect*. Yet, this cannot account for our findings, as the short run effects were present at both low and high levels of test scores, and the test score distribution is close to a normal distribution in Figure 2f, which would not be the case if test scores were restrained in the top tail.

4.2 Effects on Other Skills

The multidimensionality of skills is well-recognized (see e.g., Almlund et al., 2011) with potential cross-complementarities being an important component of skill formation (Cunha et al., 2010). In addition, it may be that the increased focus on language skills in the intervention crowded out other types of investments in the preschool setting.

To assess whether this influences the effects of the intervention, we next turn to effects of the language intervention on math test scores and well-being measured in school. Table 5 presents the results for math test scores estimated from Eq. (2). As was the case for language skills, there are large SES gaps at baseline for math skills: children from college educated parents score over 60% of a standard deviation higher than children with low educated parents. The treatment-control differences also resemble the pattern for language test scores: treated children with low educated parents experience a significant increase in test scores of 0.18 of a standard deviation, while the math test scores for treated children with college educated parents are no different than those in the control group.

For measures of well-being in columns 2-4 of Table 5, we see that the effects generally follow the same pattern as for language and math test scores. Children in the treatment group with low educated parents have significantly higher well-being and lower socio-emotional distress compared to children in the control group.

While the treatment — the intervention fostering language and pre-literacy development — only

Table 5: Long-run Differences in Math and Social Skills by Treatment Status

	Math	Well-being Survey		
	(1) Math (Grade 3)	(2) General well-being	(3) Social skills	(4) Socio-em. distress
Treated × High school/less	0.181**	0.173*	−0.053	−0.325***
	(0.086)	(0.102)	(0.092)	(0.120)
Treated × College/more	−0.019	−0.071	0.056	−0.073
	(0.086)	(0.083)	(0.082)	(0.069)
College/more	0.624***	0.121	0.060	−0.231**
	(0.068)	(0.083)	(0.095)	(0.103)
Constant	−2.700**	−0.310	−3.196***	1.826*
	(1.022)	(1.131)	(0.930)	(0.943)
Covariates	X	X	X	X
Observations	1,635	1,339	1,339	1,339

Note: Regressing children's outcomes on treatment status and parental educational attainment. Main analysis sample excluding children with an immigration background. Standard errors (in parentheses) clustered at institution level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Math test scores come from national standardized tests at grade 3 (a set-up entirely comparable to the reading test scores we use in this paper). The Well-being survey gives self-reported items by the students in grade 2. The three scales are factors resulting from an exploratory factor analysis with all items except 2 that asses the physical environment at school. Well-being and social skills are coded such that higher scores are better outcomes. A higher Socio-emotional distress score indicates a worse outcome. Other covariates included are as in Table 4, not parental years of education since we control for education levels. Full results in Table A.6.

focused on language skills and not on other competencies such as math and social skills, the results in Table 5 show no sign of crowding out of other skills. Instead, we interpret the treatment-control differences as evidence of cross-productivity from the intervention's initial effects on language skills.

5 Mechanisms

We now investigate the divergent treatment effects by grade 2 further, focusing on the role of parental investments and school quality, and the consequences for inequality in skill formation. Section 5.1 first introduces potential sources of heterogeneity in treatment effects based on a technology of skill formation (Heckman and Cunha, 2007). Section 5.2 then provides empirical evidence of changes in parents' investment behavior following the treatment, and Section 5.3 shows the mediating role of school quality for the effects of the treatment on test scores in grade 2. Finally, Section 5.4 investigates the rationale behind parents' responses and Section 5.5 discusses the implications of our findings for inequality in skill formation during childhood.

5.1 Technology of Skill Formation

We consider a standard production function for end-of-period skills θ_t that includes self-productivity from θ_{t-1} and investments I_t:[23]

$$\theta_t = j(\theta_{t-1}, I_t) \tag{3}$$

Our departure from Heckman and Cunha (2007) is to allow investments themselves to be a function of parental direct investments and investments taking place via (public) institutional settings. Parents shape parental investments $p(x_t, \theta_P)$ through direct time investments x_t, of which the efficacy depends on their own skills θ_P. The public's investments depend on an exogenously set level $\underline{G_t}$ as well as local variations. Parents can buy into neighborhoods or higher quality institutions with their wages, which are also a function of their skills, such that the public investment side via institutions depends on $g(\theta_P, \underline{G_t})$:

$$I_t = m\left(p(x_t, \theta_P), g(\theta_P, \underline{G_t})\right). \tag{4}$$

The RCT increased investments in $t = 1$ exogenously, meaning this investment was outside of parents' optimization. The result was increased child skills at the end of the period, θ_1. We assume that parents observe their child's skills at the end of the period. Parents would therefore see higher child skills after treatment, regardless of knowledge of the interventions' activities or treatment assignment. Parents then adapt their behavior in reaction to the child's development. This perspective focuses on child-driven effects, which is very frequently adopted in the literature (such as described by Bornstein et al., 2020).[24] It reflects the transactional nature of the relation between parents' and children's behaviors, and their dynamic relationship. In some economic models (such as Del Boca et al., 2019), children are even explicitly included as actors. We do not expand our model this far, but allow parents to react to their children's observed skills.

The *long-run* treatment effect of the RCT (until the end of period 2) works through this exoge-

[23] To make the model implications tractable, we consider one dimension of child skills and one dimension of parental investments. Appendix Section B.1 presents the model in detail.

[24] Child-driven effects have also been demonstrated in numerous studies relating to specific child behaviors, such as externalizing behavior. See the meta-analysis by Yan (2021), who study the strength of child-driven effects in eliciting changes in parents' psychological stress and parenting practices.

nous increase in θ_1. Formally, the long-run treatment effect is given by the **total derivative** of θ_2 with respect to θ_1:

$$\frac{d\theta_2}{d\theta_1} = \underbrace{\frac{\partial\theta_2}{\partial\theta_1}}_{\substack{\text{Self-productivity} \\ \text{of increased } \theta_1}} + \underbrace{\frac{\partial\theta_2}{\partial x_2}}_{\substack{\text{Productivity} \\ \text{of parents'} \\ \text{time investments}}} \cdot \underbrace{\frac{\partial x_2}{\partial\theta_1}}_{\substack{\text{Re-optimization} \\ \text{of parents'} \\ \text{investments} \\ \text{after intervention}}} \tag{5}$$

The heterogeneous treatment effects on grade 2 test scores viewed through the lens of this technology of skill formation suggest three potential channels: differences in self-productivity, differences in productivity of parents' investments, or differences in parents' responses to the intervention. The following subsections will provide evidence on the role of each of the three potential mechanisms.

5.2 Parents' Responses

We first investigate empirically whether the intervention and subsequent short run effects on children's skills have induced parents to adjust their investments ($\partial x_2/\partial\theta_1$).

Table 6 presents results for parents' investments based on Eq. (2). We consider two types of investments (factors); reading investments (based on items such as "How many times last week has your child been read to (or read with) at home?") and non-cognitive investments (based on items such as "When I play or read with my child, it is important to finish before we stop.").

There are substantial differences in reading investments across parents' background in the control group: college educated parents invest more in their children on average compared to low educated parents. Turning to the treatment-control differences, column 1 shows that low educated parents provide higher levels of language investments in the treatment group than in the control group. In contrast, we do not find any differences for college educated parents. Furthermore, column 2 shows that low educated parents in the treatment group also increase non-cognitive investments whereas college educated parents in the treatment group provide fewer non-cognitive investments compared to their counterparts in the control group.

Thus, the intervention and the subsequent skill improvements appear to have induced parental behavior to change, as also found in Fredriksson et al. (2016); Gelber and Isen (2013); Pop-Eleches and Urquiola (2013) among others. However, the changes in parental investment behavior vary

Table 6: Treatment–Control Differences in Parental Investments

	(1) Reading Inv.	(2) Non-cog Inv.
Treated × High school/less	0.183*	0.195*
	(0.106)	(0.101)
Treated × College/more	0.100	−0.145*
	(0.083)	(0.079)
College/more	0.250***	0.037
	(0.088)	(0.100)
Constant	0.328	2.542***
	(1.136)	(0.903)
Covariates	X	X
Observations	1,338	1,338

Note: Regression of parental investments (factors from survey) on child's treatment status interacted with parents' own education. Standard errors (in parentheses) clustered at the preschool level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Other covariates as in Table 4. Full results in Table A.7.

across the educational background: less educated parents invest more in their children's language and non-cognitive skills after the treatment, while college educated parents in the treatment group provide lower levels of investments in non-cognitive skills than college educated parents in the control group.

5.3 The Role of Differences in School Quality

As illustrated in our technology of skill formation, parents do not invest in a vacuum, because total investments in child skills depend both on investments made at home $p()$ and in an institutional setting $g()$.

As a first indication of the role of school quality, Table 7 compares the effects of the intervention by parental education (as already presented in Table 4) and by school quality based on Eq. (2) (where the interaction terms with parental education are replaced with interaction terms with school quality). The table shows that the heterogeneity is very similar across the two group definitions (even though the treatment effect in the low-quality group has a p-value of 0.10). The remarkable similarity of the point estimates begs two questions: i) are the differential treatment effects by parental education simply a proxy for differences across schools, or an interaction between the two characteristics? and ii) does school quality induce differences in the self-productivity of skills $\partial\theta_2/\partial\theta_1$, or in the productivity of parents' investments $\partial x_2/\partial\theta_1$?

Table 7: Heterogeneous Treatment Effects on Reading Skills Grade 2, by Parent Education and School Quality

	(1)	(2)
Treated × High school/less	0.232**	
	(0.021)	
Treated × College/more	−0.032	
	(0.681)	
College/more	0.480***	
	(0.000)	
Treated × School quality low		0.266
		(0.104)
Treated × School quality high		−0.015
		(0.850)
School quality high		0.336**
		(0.040)
Covariates	X	X
Observations	1,841	1,841

Note: The table shows treatment effects on reading skills in grade 2 by parental education and school quality. P-values (in parentheses), based on standard errors clustered at the preschool level. $^{*}p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$. Other covariates included are as in Table 4, and the negligible differences to those results are due to the sample restriction on available school quality here. Table 4 Full results in Table A.9.

We thus investigate whether the treatment effects differ *across school quality for a given level of parental education*. We estimate this by interacting the treatment dummy T_i with parents' education — college vs. high school (HS) — and quality of the school — low quality vs. high quality (lowQ and highQ):

$$y_i = \alpha_0 + \alpha_1 \text{HS}_i \times \text{highQ}_i + \alpha_2 \text{college}_i \times \text{lowQ}_i + \alpha_3 \text{college}_i \times \text{highQ}_i + \quad (6)$$

$$\beta_0 T_i \cdot \text{HS}_i \times \text{lowQ}_i + \beta_1 T_i \cdot \text{HS}_i \times \text{highQ}_i + \beta_2 T_i \cdot \text{college}_i \times \text{lowQ}_i + \beta_3 T_i \cdot \text{college}_i \times \text{highQ}_i + \epsilon_i$$

The estimates of α_0 to α_3 will capture the control group averages for each group, and the estimates of β_0 to β_3 will capture the average treatment-control differences for each group.

Table 8 shows the estimated treatment-control differences for children's reading test scores (column 1) and parents' investments (language and non-cognitive in columns 2 and 3, respectively), by the quality of the school and parental education level. The positive effects on reading are concentrated among children with *low educated parents in low quality schools.* For these children, the grade 2 test scores in the treatment group are 0.33 of a standard deviation higher than in the control group. In contrast, there are no significant effects for children with low educated parents *in high quality schools* or children with college educated parents (regardless of school quality). The treatment-control difference of -0.085 for children with college educated parents in high quality schools is statistically significantly lower than the treatment-control difference for the two groups with low educated parents (0.33 and 0.176), suggesting that school quality also plays a mediating role for these groups, albeit estimated with less precision (Table A.11 presents the full set of pairwise hypothesis tests).

Furthermore, the results from column 1 in Table 8 illustrate that differences in self-productivity are an unlikely explanation of the heterogenous treatment effects ($\partial\theta_2/\partial\theta_1$ from Eq. (5)). If self-productivity was the main mechanism behind the different effects for children with low educated parents in low quality schools vs. children with college educated parents in high quality schools, we should i) see similar effects for children with low educated parents irrespective of school quality because of the similar baseline reading test scores observed for the control group (see the insignificant estimate of 0.146 for $High\ school/less \times School\ quality\ high$ in column 1) and ii) different treatment effects on language test scores for children with low educated parents in high quality schools vs. children with college educated parents in low quality schools (compare 0.176 for $Treated \times High\ school/less \times School\ quality\ high$ to 0.190 for $Treated \times College/more \times School\ quality\ low$ in column 1) — but we do not observe such a pattern.

These heterogeneous effects on reading skills in grade 2 are robust to alternative measures. First, when we use the quality of a child's predicted school (on the basis of where the majority of classmates will go after preschool) instead of the realized school as throughout the main text, the results are not statistically distinguishable (see Table A.13). Second, when we use the alternative measure of school quality that predicts average attainment on the basis of a rich set of teacher characteristics at each school, the positive treatment effects on reading skills remain concentrated in the group of children with less educated parents attending worse schools (Table A.12). In fact, the treatment effects here turn significantly negative for children of highly educated parents attending the best schools.

To investigate the potential mechanisms underlying the heterogenous treatment effects further, we turn to treatment-control differences in parental investments. Here, columns 2 and 3 in Table 8 show that low educated parents whose children attend low quality schools increase investments

Table 8: Treatment Effects by Parental Education and School Quality

	(1) Reading Test Grade 2	(2) Reading Investments	(3) Non-cognitive Investments
Treated × High school/less × School quality low	0.332***	0.497***	0.282
	(0.124)	(0.184)	(0.177)
Treated × High school/less × School quality high	0.176	0.043	0.061
	(0.124)	(0.140)	(0.132)
Treated × College/more × School quality low	0.190	0.019	0.028
	(0.254)	(0.145)	(0.175)
Treated × College/more × School quality high	−0.085	0.075	−0.234**
	(0.071)	(0.107)	(0.101)
High school/less × School quality low	0.000	0.000	0.000
	(.)	(.)	(.)
High school/less × School quality high	0.146	0.277*	0.270
	(0.140)	(0.148)	(0.185)
College/more × School quality low	0.268	0.414***	0.187
	(0.161)	(0.147)	(0.184)
College/more × School quality high	0.649***	0.462***	0.180
	(0.131)	(0.148)	(0.164)
Covariates	X	X	X
Observations	1,841	1,129	1,129

Note: "School quality low" is defined as school being in the bottom quartile of school quality in terms of average test scores in the RCT sample, which is below the 40th percentile of the national distribution. The reverse for "School quality high." Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. For full results, see Table A.10. Table A.11 contrasts all coefficients on the treatment-interactions in pairwise hypothesis tests.

substantially for the treatment group relative to the control group. The effect of 0.497 is large. It is larger than the gap across school quality within low-educated parents, where parents of children attending higher-quality schools invest 0.277 more than when children attend low-quality schools. This means that treated parents in the most disadvantaged group (low education and low-quality schools) more than than compensate for their otherwise low investments.[25] The gap in language investments among low educated parents across school quality is more than closed due to the treatment response. Moreover, college educated parents in schools with high quality *reduce* their non-cognitive investments significantly (column 3).

In sum, the treatment-control differences in parents' investments strongly coincide with the effects we observe for test scores. For children with low educated parents in low-quality schools, we observe positive and significant treatment-control differences in both reading test scores and parents' investments. Among children with low educated parents in high-quality schools and children with college educated parents in low-quality schools, we observe positive but statistically insignificant treatment-control differences in reading test scores and parents' investments. Finally, for children with college educated parents in high-quality schools, we observe significantly lower treatment-control differences than for children with low educated parents in both low- and high-quality schools. Moreover, we also observe a significant reduction in non-cognitive investments for college educated parents with children in high-quality schools.[26]

5.4 Explaining Parents' Responses

While the results from Section 5.3 show that the differences in treatment effects on test scores likely stem from parents' responses to the intervention, it is not clear why some parents respond by increasing investments in their children while others respond by reducing investments.

As a lever to illustrate parents' trade-offs when choosing investment levels, Appendix Section B.1 incorporates the technology of skill formation from Section 5.1 in a model where parents' choose investment levels x, leisure l, and labor supply h to maximize their utility, which depends on

[25] Column 2 shows a 0.277 difference between low educated parents' language investments across school quality in the *control* group. The treatment-control differences for low educated parents in high and low quality schools, respectively, imply that this gap is closed in the treatment group ($0.497 - 0.043 = 0.454$).

[26] The absence of differences between treatment effects for children with low educated parents in high quality schools (second row, Table 8) and children with college educated parents in low quality schools (third row, Table 8) suggest that differences in the productivity of parents' investments do not play an important role behind the heterogeneous treatment effects.

consumption, leisure, and children's skills.[27] Parents set investments and leisure such that the marginal utility of investments equals the marginal utility of leisure:

$$\beta \frac{\partial V}{\partial \theta_2} \frac{\partial \theta_2(\theta_1, x_2, \theta_P, \underline{G}_2)}{\partial x_2} = \frac{\partial u_2}{\partial l_2} \tag{7}$$

where V is parents' value from children's long run skills (θ_2), x_2 are parents' investments in child skills, and l_2 parents' leisure. β is the discount factor.

This simple but general model does not yield strong predictions for how parents change investments after the initial improvement in their child's skills from the intervention. The reactions depend theoretically mostly on assumptions about the elasticity of substitution between skills, parental investments, and investments in schools. In the literature, empirical results on the elasticity of substitution between child skills and parental investments are mixed: Findings from Bonesrønning (2004); Chang et al. (2022); Gelber and Isen (2013) suggest that parental and institutional investments are complements, while findings from Das et al. (2013); Fredriksson et al. (2016); Greaves et al. (2023); Houtenville and Conway (2008); Pop-Eleches and Urquiola (2013) suggest that they are substitutes.

Our model assumes only that skills are not detrimental (not negative self-productivity), but notes that parents are generally more likely to reduce their investments — except if the exogenous increase in child skills raised the productivity of parental investments.

As an alternative approach, we provide empirical evidence of the initial differences in parents' choices and potential constraints, and relate these to parents' observed investment responses. Column 1 in Table 9 shows the differences in children's time spent in non-parental care (school and after-school care) by parents' education level and school quality, in the control group. Children's time in non-parental care decreases as parents' education levels increase.[28] Reduced time *not* with parents can be interpreted as increased time *with* parents. This would further corroborate the previous findings from Sections 5.2 and 5.3 that college educated parents invest more in their children than low educated parents.

Parents' hours of work are also positively associated with their education level and school quality

[27] Formally, parents' utility U in the two-period model is $U(c_1, c_2, l_1, l_2, \theta_2) = u(c_1, l_1) + \beta u(c_2, l_2) + \beta^2 V(\theta_2)$.

[28] In the Danish context, after-school care is non-academic supervised play time—therefore, the length of the school day including this after-school activities does *not* correspond to academic investments.

Table 9: Work / Non-parental Care Times in Control Group

	(1) Hrs School	(2) Hrs Work
High school/less × School quality high	−0.071	1.538
	(0.177)	(1.277)
College/more × School quality low	0.011	1.114
	(0.214)	(2.227)
College/more × School quality high	−0.315*	4.300***
	(0.169)	(1.092)
Constant	7.829***	31.284***
	(0.140)	(1.033)
Observations	571	571

Note: Sample for this table: only children in the control group, based on responses by parents in the follow-up survey. "Hrs in Non-parental Care" refer to the child's length of the standard school day including non-academic after-school care, measured by the typical time the child is picked up minus the typical time parents drop the child off. "Hrs Work" are typical weekly work hours as reported by parents. No other covariates included. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

(column 2 in Table 9). Thus, although we do not have precise measures of parents' time use more broadly, it is clear from the patterns for the control group in Table 9 (children's hours spent in school and parents' hours of work) and Table 8 (children's test scores and parents' investments) that highly educated parents already invest a lot, more than less-educated parents, and work longer hours; *especially* so when their children are attending high-quality schools. Therefore, the marginal value of changes in investments and leisure are bound to differ substantially across different levels of parents' education and school quality. College educated parents with children in high quality schools have most to gain from trading investments for leisure (assuming convex preferences). In contrast, low educated parents with children in low quality schools have most to gain in terms of supplementing the intervention with increased investments in skills.

5.5 Implications for Inequality in Children's Skills

To quantify how differences in parental responses and effects on language test scores from the treatment affected socio-economic inequalities, Fig. 3a and b present the predicted language test scores in preschool (3a) and grade 2 (3b) across parents' average years of schooling and school quality based on the regression:

$$y_i = \alpha_0 + \alpha_1 \text{SES}_i + \alpha_2 \text{SES}_i^2 + \beta_1 T_i \times \text{SES}_i + \beta_2 T_i \times \text{SES}_i^2 + \epsilon_i \tag{8}$$

where y_i is children's test scores, T_i is treatment status, and SES_i and SES_i^2 are indicators of Socio-Economic-Status (parents' years of schooling or school quality) linearly and squared, respectively (we include the squared term to allow for potential nonlinearities).

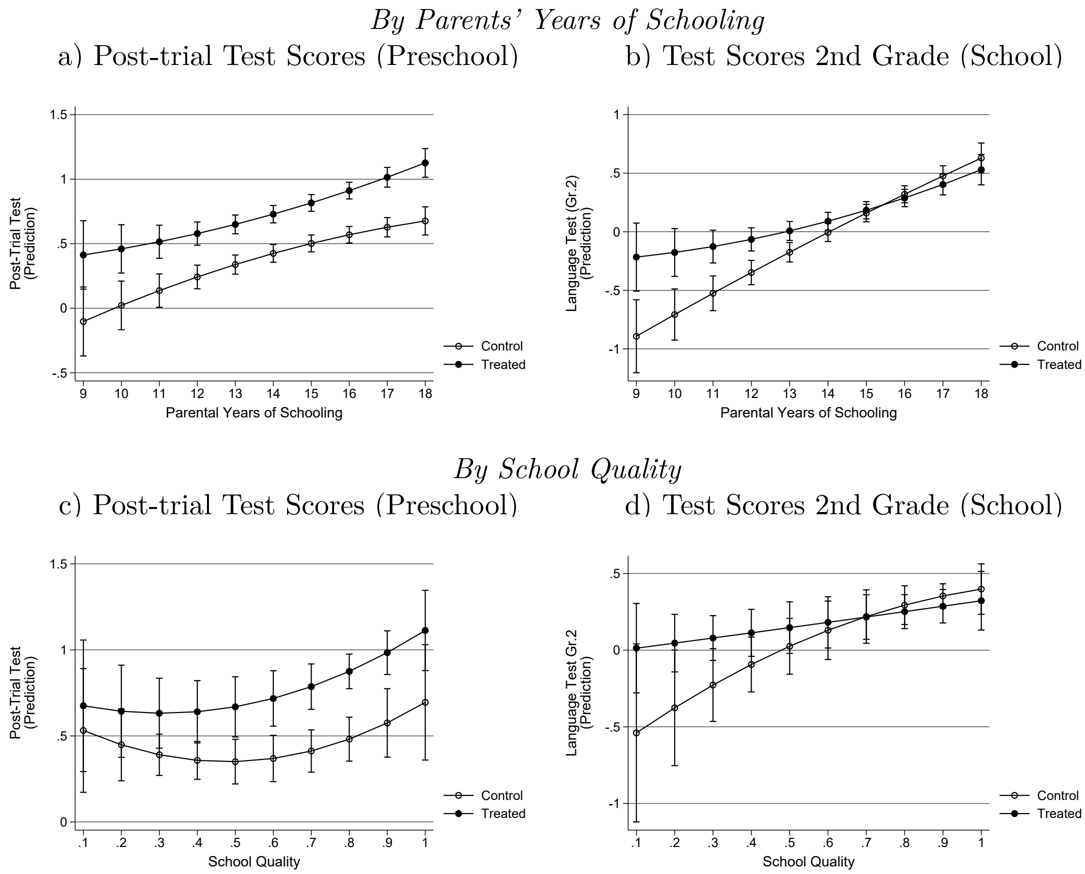
The figures illustrate the stark contrast between the treatment effects in preschool and school ages. In Fig. 3a, the differences in children's test scores measured in preschool across parents' education levels are large; going from 9 to 18 years of schooling is associated with an average test score increase of 0.7 standard deviations. Moreover, the gradients are parallel in the control and treatment group, reflecting the positive treatment effects observed for all groups. Thus, there is initially no indication of the intervention with increased focus on language development in preschool would level the playing field across parental background.

Inequality in language skills widens between preschool and grade 2 of elementary school. As Figure 3b demonstrates, there is a steeper gradient in children's test scores by parents' education levels in grade 2 relative to preschool in the control group. Going from 9 to 18 years of schooling is associated with an average reading improvement of 1.3 standard deviations. In the treatment group, the association between children's test scores and parents education is substantially attenuated; going from 9 to 18 years of schooling is associated with an average increase in child test scores of 0.6-0.7 standard deviation, which is 50% less than the corresponding gradient in the control group.

Figures 3c and d present similar results by school quality based on Eq. (8) where we replace parents' years of schooling with school quality in the estimation. The results show a strong reduction in test score differences in the treatment group across school quality compared to the control group: Going from the lowest to the highest quality school, reading test scores in grade 2 increase by almost 1 standard deviation in the control group, whereas the similar difference in the treatment group is only around 0.4 of a standard deviation.

To illustrate the convergence in language test scores across both parents' education and school quality for children in the treatment group, Fig. 4 shows cumulative distributions of reading test scores in grade 2 by parents' education levels and school quality for the control group (4a) and

Figure 3: Association Between Child Language Test Scores, and Parent Education and School Quality by Treatment Status

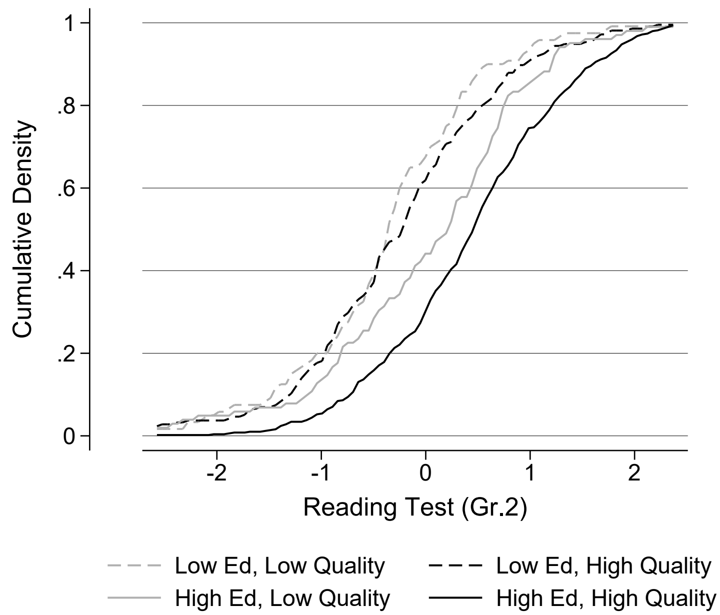


Note: The figure shows the predicted post-trial and grade 2 language test scores from regressions where a linear and quadratic term for parental highest years of education (a and b) / school quality (c and d) have been interacted with a treatment indicator, with 90% confidence intervals. Test scores are standardized to mean of 0 and standard deviation of 1. Standard errors are clustered at the preschool level.

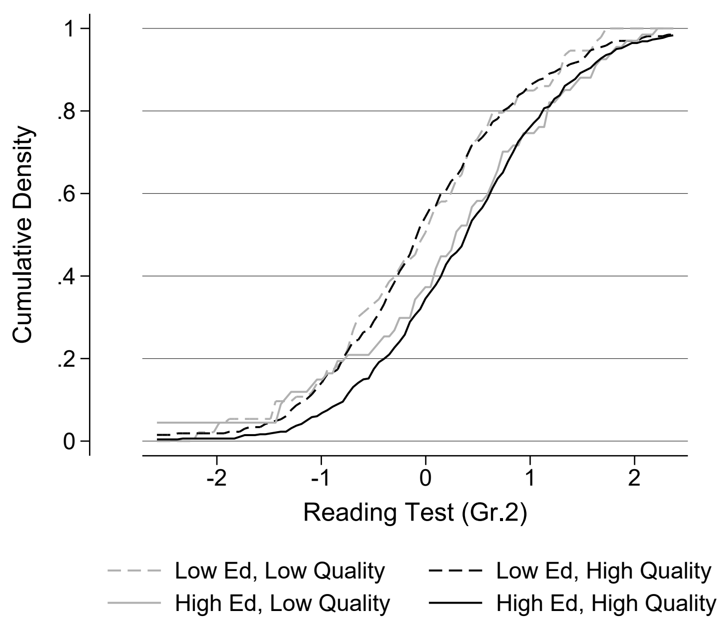
treatment group (4b). In the control group, the figure shows clear differences by parents' education and school quality across the entire test score distribution. However, in the treatment group the differences are more muted in all parts of the distribution and in some test score ranges the differences between low and high quality schools are almost eliminated.

Figure 4: Reading Test Scores (Grade 2), by Treatment Status and Parental Education/School Quality

a) Control Group



b) Treatment Group



Note: The figures show the cumulative distributions of grade 2 reading test scores by parents' education and school quality separately for the control group (a) and the treatment group (b).

6 Conclusion

In this paper, we study how public and parental investments interact in the formation of children's skills, from preschool to elementary school. We do so by analyzing a randomized controlled trial in Denmark, which exogenously increased the quality of regular public preschools. By linking the children from this intervention to full population register data with information on later school quality and test scores, we find that increasing the quality of universal childcare reduces skill gaps between children from advantaged and disadvantaged backgrounds. The intervention had large initial positive effects on language skills at ages 4 to 5 for children across *all* backgrounds. For children with low educated parents, test scores remain 0.22 standard deviations higher in the treatment group than in the control group at ages 8 to 9 (4 to 5 years after the intervention), while there is a complete fade-out for children with college educated parents. We also find evidence of cross-productivity from the intervention, since following improved language skills, children's self-reported well-being in grade 2 and math test scores measured in grade 3 are significantly different in the treatment vs. control group.

We show that the heterogenous treatment effects likely result from parents' responses to the intervention (and not a result of other potential factors such as ceiling effects) in combination with the quality of the schools children attend. Children with low educated parents in low quality schools (measured via the composition of peers and teachers in each school) drive the positive treatment effects. Parents invest more in both language and non-cognitive skills in this group. In contrast, the complete fade-out for children with college educated parents is driven by those who attend high quality schools, where we observe reductions in some dimensions of parents investments. Our findings thereby underscore the importance of complementing early childhood investments with a strong focus on children's home and out-of-home environments in the subsequent years.

References

Abdulkadiroğlu, A., P. A. Pathak, and C. R. Walters (2018). Free to Choose: Can School Choice Reduce Student Achievement. *American Economic Journal: Applied Economics 10*(1), 175–206.

Agostinelli, F., M. Saharkhiz, and M. J. Wiswall (2024). Home and school in the development of children. *International Economic Review* (forthcoming).

Almlund, M., A. Duckworth, J. J. Heckman, and K. T (2011). *Personality psychology and economics*, Volume 4 of *Handbook of the Economics of Education*. Elsevier.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–172.

Andersen, S. C., M. Gensowski, S. G. Ludeke, and O. P. John (2020). A stable relationship between personality and academic performance from childhood through adolescence. An original study and replication in hundred-thousand-person samples. *Journal of Personality 88*(5), 925–939.

Ansari, A. and R. C. Pianta (2018). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review 86*, 120–127.

Attanasio, O. P., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia. *American Economic Review*, forthcoming.

Bailey, M. J., B. D. Timpe, and S. Sun (2020). Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency. *NBER Working Paper* (28268).

Baker, M., J. Gruber, and K. Milligan (2008). Universal Child Care, Maternal Labor Supply, and Family Well-Being. *Journal of Political Economy 116*(4), 709–745.

Bayer, P., F. Ferreira, and R. McMillan (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bénabou, R. (1996). Equity and Efficiency in Human Capital Investment: The Local Connection. *The Review of Economic Studies 63*(2), 237–264.

Bernal, R. and M. P. Keane (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics 156*(1), 164–189.

Beuchert, L. V. and A. B. Nandrup (2018). The Danish National Tests at a Glance. *Nationaloekonomisk Tidsskrift 1*, 1–35.

Bjerre-Nielsen, A. and M. H. Gandil (2020). Defying Attendance Boundary Policies and the Limits to Combating School Segregation. *Working Paper, University of Copenhagen*, 1–48.

Blanden, J., M. Doepke, and J. Stuhler (2023). Educational Inequality. In E. A. Hanushek, L. Woessmann, and S. J. Machin (Eds.), *Handbook of the Economics of Education* (1 ed.)., Chapter 6. Amsterdam: North Holland.

Bleses, D., A. Højen, P. S. Dale, L. M. Justice, L. Dybdal, S. Piasta, J. Markussen-Brown, L. Kjærbæk, and E. F. Haghish (2018). Effective language and literacy instruction: Evaluating the importance of scripting and group size components. *Early Childhood Research Quarterly 42*(October 2017), 256–269.

Bonesrønning, H. (2004). The determinants of parental effort in education production: Do parents respond to changes in class size? *Economics of Education Review 23*(1), 1–9.

Bornstein, M. H., D. L. Putnick, and G. Esposito (2020, October). Skill–experience transactions across development: Bidirectional relations between child core language and the child's home learning environment. *Developmental Psychology 56*(10), 1842–1854.

Burgess, S., E. Greaves, A. Vignoles, and D. Wilson (2015). What Parents Want: School Preferences and School Choice. *Economic Journal 125*(587), 1262–1289.

Carneiro, P. and R. Ginja (2014). Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start. *American Economic Journal: Economic Policy 6*(4), 135–173.

Cascio, E. U. (2021). Early Childhood Education in the United States: What, When, Where, Who, How, and Why. In B. McCall (Ed.), *Routledge Handbook of the Economics of Education*, Chapter 2, pp. 30–72. New York.

Cascio, E. U. and D. W. Schanzenbach (2013). The Impacts of Expanding Access to High-Quality Preschool Education. *Brookings Papers on Economic Activity, Economic Studies Program 47*(2), 127–192.

Chang, S., D. A. Cobb-Clark, and N. Salamanca (2022). Parents' Responses to Teacher Qualifications. *Journal of Economic Behavior & Organization 197*(13065), 419–446.

Chaparro, J., A. Sojourner, and M. J. Wiswall (2020). Early Childhood Care and Cognitive Development. *NBER Working Paper* (26813).

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Cornelissen, T., C. Durstmann, U. Schönberg, and A. Raute (2018). Who Benenfits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy 126*(6), 2356–2409.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica 78*(3), 883–931.

Dahl, M., C. S. H. Anette Eklund Hansen, and J. E. Kristensen (2015). *Kamp og status – de lange linjer i bornehaveinstitutionens of paedagogprofessionen historie 1820 til 2015.* Paedagogprofessionens historie og aktualitet.

Danish Ministry for Children and Social Affairs (2018). Bekendtgørelse om pædagogiske mål og indhold i seks læreplanstemaer. Technical report.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2013). School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics 5*(2), 29–57.

Datta Gupta, N. and M. Simonsen (2010a). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics 94*(1-2), 30–43.

Datta Gupta, N. and M. Simonsen (2010b). Non-Cognitive Child Outcomes and Universal High Quality Child Care. *Journal of Public Economics 94*(1–2), 30–43.

Datta Gupta, N. and M. Simonsen (2012). The Effects of Type of Non-Parental Child Care on Pre-Teen Skills and Risky Behavior. *Economics Letters 116*(3), 622–625.

Del Boca, D., C. J. Flinn, E. Verriest, and M. J. Wiswall (2019). Actors in the Child Development Process. *NBER Working Paper 25596*.

Del Boca, D., C. J. Flinn, and M. Wiswall (2014). Household Choices and Child Development. *Review of Economic Studies 81*(1), 137–185.

Eshaghnia, S., J. J. Heckman, and G. Razavi (2023). Pricing Neighborhoods. *NBER Working Paper* (31371).

Esping-Andersen, G., I. Garfinkel, W.-J. Han, K. Magnuson, S. Wagner, and J. Waldfogel (2012). Child Care and School Performance in Denmark and the United States. *Child Youth Services Review 34*(3), 576–589.

European Commission/EACEA/Eurydice (2019). *Key Data on Early Childhood Education and Care in Europe – 2019 Edition*. Euridice Report. Luxembourg: Publications Office of the European Union.

Fort, M., A. Ichino, and G. Zanella (2020). Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families. *Journal of Political Economy 128*(1), 158–205.

Fredriksson, P., B. Öckert, and H. Oosterbeek (2016). Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule. *Journal of Human Resources 51*(4), 832–868.

Gelber, A. M. and A. Isen (2013). Children's schooling and parents' behavior: Evidence from the Head Start Impact Study. *Journal of Public Economics 101*, 25–38.

Gormley, W. J., D. Phillips, and T. Gayer (2008). Preschool Programs Can Boost School Readiness. *Science 320*(320), 1723–1724.

Greaves, E., I. Hussain, B. Rabe, and I. Rasul (2023). Parental responses to information about school quality: Evidence from linked survey and administrative data. *The Economic Journal 133*(654), 2334—-2402.

Havnes, T. and M. Mogstad (2011). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy 3*(2), 97–129.

Havnes, T. and M. Mogstad (2015). Is Universal Child Care Leveling the Playing Field? *Journal of Public Economics 127*, 100–114.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science 312*(5782), 1900–1902.

Heckman, J. J. and F. Cunha (2007). The Technology of Skill Formation. *American Economic Review 97*(2), 31–47.

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review 103*(6), 2052–2086.

Højen, A., D. Bleses, and P. S. Dale (2022, August). *Language Assessments for Preschool Children: Validity and Reliability of Two New Instruments Administered by Childcare Educators* (1 ed.). Cambridge University Press.

Houtenville, A. J. and K. S. Conway (2008). Parental Effort, School Resources, and Student Achievement. *The Journal of Human Resources 43*(2), 437–453.

Justice, L. M., J. A. Logan, J. N. Kaderavek, and J. M. Dynia (2015). Print-Focused Read-Alouds in Early Childhood Special Education Programs. *Exceptional Children 81*(3), 292–311.

Justice, L. M. and A. S. McGinty (2012). Read It Again-PreK! A Preschool Curricular Supplement to Promote Language and Literacy Foundation. Technical report, Crane Center for Early Childhood Research and Policy.

Kraft, M. A., D. Blazar, and D. Hogan (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research 88*(4), 547–588.

Kragh-Müller, G. (2017). The key characteristics of Danish/Nordic child care culture. In *Nordic social pedagogical approach to early years*, pp. 3–23. Springer.

Landersø, R. and J. J. Heckman (2016). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S. *Scandinavian Journal of Economics 119*(1), 178–230.

Markussen-Brown, J., C. B. Juhl, S. B. Piasta, D. Bleses, A. Højen, and J. L. (2017). The Effects of Language- and Literacy-Focused Professional Development on Early Educators and Children: A Best-Evidence Meta-Analysis. *Early Childhood Research Quarterly 38*, 97–115.

Mogstad, M. and G. Torsvik (2023). Family background, neighborhoods, and intergenerational mobility. In S. Lundberg and A. Voena (Eds.), *Handbook of the Economics of the Family*, Volume 1, Chapter 6, pp. 327–387. North-Holland.

Nicoletti, C. and B. Rabe (2014). School Inputs and Skills: Complementarity and Self-Productivity. *IZA Discussion Paper* (8693).

OECD (2023). *Education at a Glance 2023*.

Pop-Eleches, C. and M. Urquiola (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review 103*(4), 1289–1324.

Sievertsen, H. H., F. Gino, and M. Piovesan (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences of the United States of America 113*(10), 2621–2624.

Slot, P. (2018). Structural characteristics and process quality in early childhood education and care: A literature review. Technical Report Working Paper No. 176, OECD.

Slot, P. L., D. Bleses, L. M. Justice, J. Markussen-Brown, and A. Højen (2018). Structural and Process Quality of Danish Preschools: Direct and Indirect Associations With Children's Growth in Language and Preliteracy Skills. *Early Education and Development 29*(4), 581–602.

Yan, N., A. A. P. P. (2021). Reconsidering the relation between parental functioning and child externalizing behaviors: A meta-analysis on child-driven effects. *Journal of Family Psychology 35*(2), 225–235.

A For Online Publication: Public and Parental Investments, and Children's Skill Formation

Table A.1: Balancing Test Across Treatment Status by Parents' Education

	(1) All	(2) Parent Edu: High school/less	(3) Parent Edu: College/more
Pre-Trial Test	0.022	0.022	0.025
	(0.020)	(0.024)	(0.022)
Born in 2007	−0.075	0.244	−0.282*
	(0.113)	(0.173)	(0.164)
Born in 2008	−0.051	0.203	−0.231
	(0.124)	(0.178)	(0.168)
Born in 2009	−0.031	0.209	−0.206
	(0.124)	(0.182)	(0.178)
Born in 2010	−0.097	0.075	−0.235
	(0.160)	(0.200)	(0.209)
Male	−0.027	−0.051	−0.011
	(0.026)	(0.034)	(0.034)
Birth weight (kg)	−0.061**	−0.101**	−0.031
	(0.027)	(0.039)	(0.037)
Gestation (wks)	0.002	−0.010	0.005
	(0.009)	(0.012)	(0.012)
Apgar score	0.018	−0.001	0.027
	(0.021)	(0.032)	(0.024)
Number of Siblings	−0.019	−0.026	−0.020
	(0.020)	(0.028)	(0.022)
Mother weight (kg)	0.001	−0.000	0.002*
	(0.001)	(0.001)	(0.001)
Mother education (yrs)	−0.004	0.006	−0.001
	(0.008)	(0.016)	(0.009)
Mother age	0.004	0.010*	−0.001
	(0.004)	(0.006)	(0.005)
Mother employed	0.037	0.053	0.001
	(0.036)	(0.048)	(0.048)
Father education (yrs)	−0.003	−0.006	0.004
	(0.008)	(0.012)	(0.011)
Father age	−0.005*	−0.003	−0.006
	(0.003)	(0.005)	(0.004)
Father employed	−0.011	−0.023	0.031
	(0.052)	(0.062)	(0.061)
Household inc(1,000 USD)	0.000	−0.000	0.000
	(0.000)	(0.001)	(0.000)
Constant	0.601	0.952	0.468
	(0.471)	(0.652)	(0.580)
Observations	2,301	836	1,465

Note: The table shows results from regressing treatment status simultaneously on all covariates. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A.2: Joint Test of Balance by Treatment Status — Different Outcome Samples

	(1) Language Test Gr.2	(2) Math Test Gr.3	(3) Well-being Survey	(4) Parent Survey
Pre-Trial Test	0.028	0.034	0.014	0.024
	(0.021)	(0.023)	(0.024)	(0.022)
Born in 2007	−0.100	−0.036	0.024	−0.103
	(0.121)	(0.137)	(0.169)	(0.245)
Born in 2008	−0.090	−0.032	0.043	−0.079
	(0.129)	(0.144)	(0.167)	(0.258)
Born in 2009	−0.091	−0.041	0.018	−0.053
	(0.128)	(0.144)	(0.169)	(0.260)
Born in 2010	−0.174	−0.150	−0.064	−0.133
	(0.160)	(0.193)	(0.194)	(0.286)
Male	-0.046^{*}	−0.036	-0.059^{*}	−0.018
	(0.027)	(0.029)	(0.033)	(0.034)
Birth weight (kg)	-0.064^{**}	-0.052^{*}	−0.032	−0.043
	(0.026)	(0.031)	(0.033)	(0.035)
Gestation (wks)	0.003	0.001	−0.001	0.002
	(0.009)	(0.009)	(0.010)	(0.012)
Apgar score	0.028	0.043^{*}	0.014	−0.001
	(0.021)	(0.024)	(0.024)	(0.028)
Number of Siblings	−0.019	−0.018	-0.042^{*}	−0.015
	(0.021)	(0.022)	(0.021)	(0.026)
Mother weight (kg)	0.002^{*}	0.002^{*}	0.002^{*}	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Mother education (yrs)	−0.001	−0.004	−0.004	−0.003
	(0.009)	(0.009)	(0.009)	(0.011)
Mother age	0.004	0.003	0.005	0.005
	(0.004)	(0.005)	(0.005)	(0.006)
Mother employed	0.020	0.027	0.015	0.005
	(0.036)	(0.039)	(0.040)	(0.048)
Father education (yrs)	−0.003	−0.005	−0.005	−0.005
	(0.008)	(0.009)	(0.010)	(0.010)
Father age	-0.005^{*}	-0.006^{*}	−0.004	-0.010^{***}
	(0.003)	(0.003)	(0.004)	(0.004)
Father employed	−0.031	−0.023	−0.020	0.012
	(0.059)	(0.063)	(0.060)	(0.058)
Household inc(1,000 USD)	−0.000	−0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.531	0.457	0.522	0.926
	(0.477)	(0.498)	(0.533)	(0.567)
Observations	1,898	1,635	1,339	1,338

Note: The table shows results from regressing treatment status simultaneously on all covariates for sub-samples with non-missing later outcomes (referenced in column headers). Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table A.3: Regressing Data Availability on Treatment Indicator

	(1)	(2)	(3)	(4)
Missing Language Score Gr.2	0.016			
	(0.060)			
Missing Math Score Gr.3		-0.002		
		(0.046)		
Missing Well-being Survey			0.042	
			(0.039)	
Missing Parent Survey				-0.035
				(0.031)
Constant	0.497^{***}	0.501^{***}	0.482^{***}	0.515^{***}
	(0.066)	(0.067)	(0.068)	(0.068)
Response Rate	0.825	0.711	0.582	0.581
Count Responses	1,898	1,635	1,339	1,338
Observations	2,301	2,301	2,301	2,301

Note: The table shows estimates from regressions of indicators for non-missing information in long-run language (column 1) and math tests (column 2), as well as for child response on the Danish well-being survey in grade 2 (column 3), and parents' survey response to our follow-up survey (column 4) on treatment status.

Table A.4: Test of Data Availability and Covariates

	(1) Language Test (Gr.2)	(2) Math (Grade 3)	(3) Well-being Survey	(4) Parent Survey
Treated × High school/less	0.013	−0.006	−0.017	0.001
	(0.030)	(0.028)	(0.039)	(0.043)
Treated × College/more	−0.025	−0.011	−0.059	0.043
	(0.041)	(0.036)	(0.046)	(0.034)
College/more	−0.025	−0.060**	0.023	0.061
	(0.030)	(0.029)	(0.040)	(0.041)
Pre-Trial Test	0.013	0.008	0.026**	0.039***
	(0.009)	(0.009)	(0.011)	(0.011)
Child age	0.018	0.019	−0.012	−0.023
	(0.017)	(0.021)	(0.024)	(0.021)
Born in 2007	−0.002	0.130	0.015	0.291**
	(0.083)	(0.104)	(0.115)	(0.134)
Born in 2008	0.005	0.145	−0.004	0.286**
	(0.085)	(0.100)	(0.112)	(0.135)
Born in 2009	−0.011	0.089	−0.058	0.262*
	(0.090)	(0.107)	(0.115)	(0.147)
Born in 2010	−0.029	−0.437***	−0.089	0.178
	(0.107)	(0.126)	(0.126)	(0.159)
Male	−0.015	0.005	−0.005	−0.025
	(0.016)	(0.015)	(0.022)	(0.020)
Birth weight (kg)	−0.009	−0.012	0.009	−0.015
	(0.019)	(0.021)	(0.027)	(0.026)
Gestation (wks)	0.003	0.003	0.002	0.006
	(0.005)	(0.006)	(0.008)	(0.008)
Apgar score	−0.008	−0.008	−0.021	0.018
	(0.015)	(0.019)	(0.019)	(0.018)
Number of Siblings	0.021**	0.009	0.026**	−0.016
	(0.009)	(0.010)	(0.012)	(0.012)
Mother weight (kg)	0.001	0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Mother age	−0.007***	−0.005**	−0.002	0.003
	(0.002)	(0.003)	(0.003)	(0.003)
Mother employed	0.008	−0.005	0.051	0.071**
	(0.022)	(0.023)	(0.032)	(0.032)
Father age	0.000	0.001	−0.005**	−0.000
	(0.002)	(0.002)	(0.002)	(0.003)
Father employed	0.063*	0.029	0.007	0.132***
	(0.033)	(0.034)	(0.032)	(0.037)
Household inc(1,000 USD)	−0.000	0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.845***	0.623*	1.037**	−0.208
	(0.269)	(0.367)	(0.408)	(0.412)
Response Rate	0.825	0.711	0.582	0.581
Count Responses	1,898	1,635	1,339	1,338
Observations	2,301	2,301	2,301	2,301

Note: The table shows linear regressions of indicators for non-missing information in long-run language (column 1) and math tests (column 2), as well as for child response on the Danish well-being survey in grade 2 (column 3), and parents' survey response to our follow-up survey (column 4) on treatment status and covariates.

Table A.5: Main Treatment Effects: Full Results with Covariates

	Pre-Trial Test		Post-Trial Test		National Test Grade 2	
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	By parent ed	Pooled	By parent ed	Pooled	By parent ed
Treated	0.084		0.309^{***}		0.055	
	(0.271)		(0.000)		(0.468)	
Treated × High school/less		0.079		0.304^{***}		0.222^{**}
		(0.412)		(0.000)		(0.023)
Treated × College/more		0.091		0.311^{***}		−0.039
		(0.311)		(0.000)		(0.606)
College/more		0.324^{***}		0.075		0.483^{***}
		(0.000)		(0.147)		(0.000)
Child age	−0.043	−0.046	-0.078^{*}	-0.079^{*}	0.033	0.035
	(0.358)	(0.329)	(0.056)	(0.051)	(0.562)	(0.545)
Born in 2007	0.242	0.185	−0.106	−0.120	0.130	0.059
	(0.261)	(0.388)	(0.430)	(0.398)	(0.559)	(0.795)
Born in 2008	−0.035	−0.091	−0.101	−0.116	0.224	0.162
	(0.875)	(0.682)	(0.467)	(0.425)	(0.313)	(0.485)
Born in 2009	−0.009	−0.069	−0.247	−0.263	0.290	0.221
	(0.966)	(0.747)	(0.163)	(0.147)	(0.255)	(0.403)
Born in 2010	0.130	0.062	−0.281	−0.301	0.349	0.281
	(0.569)	(0.785)	(0.170)	(0.147)	(0.239)	(0.360)
Male	−0.017	−0.020	0.009	0.009	-0.167^{***}	-0.169^{***}
	(0.656)	(0.604)	(0.776)	(0.779)	(0.000)	(0.000)
Birth weight (kg)	0.074	0.087^{*}	0.074^{**}	0.078^{**}	0.099^{**}	0.116^{**}
	(0.133)	(0.090)	(0.044)	(0.037)	(0.032)	(0.011)
Gestation (wks)	0.013	0.012	-0.023^{*}	-0.024^{*}	−0.002	−0.002
	(0.460)	(0.520)	(0.091)	(0.087)	(0.894)	(0.930)
Apgar score	−0.013	−0.016	0.021	0.020	−0.032	−0.028
	(0.649)	(0.571)	(0.630)	(0.640)	(0.189)	(0.267)
Number of Siblings	-0.088^{***}	-0.092^{***}	-0.051^{**}	-0.053^{**}	0.032	0.031
	(0.002)	(0.001)	(0.040)	(0.035)	(0.215)	(0.219)
Mother weight (kg)	−0.001	−0.001	−0.000	−0.000	−0.002	−0.002
	(0.552)	(0.344)	(0.832)	(0.691)	(0.327)	(0.213)
Mother education (yrs)	0.059^{***}		0.020^{**}		0.058^{***}	
	(0.000)		(0.021)		(0.000)	
Mother age	0.006	0.010^{*}	0.002	0.003	0.015^{**}	0.018^{***}
	(0.276)	(0.089)	(0.745)	(0.541)	(0.016)	(0.009)
Mother employed	−0.007	0.031	0.048	0.064	0.058	0.088
	(0.915)	(0.661)	(0.316)	(0.174)	(0.407)	(0.210)
Father education (yrs)	0.038^{***}		0.008		0.041^{***}	
	(0.002)		(0.412)		(0.000)	
Father age	−0.004	−0.005	−0.002	−0.002	0.001	−0.001
	(0.418)	(0.297)	(0.634)	(0.595)	(0.901)	(0.879)
Father employed	0.043	0.058	−0.052	−0.048	0.046	0.076
	(0.500)	(0.346)	(0.366)	(0.407)	(0.536)	(0.311)
Household inc(1,000 USD)	0.000	0.001^{**}	0.000	0.000	−0.000	−0.000
	(0.254)	(0.034)	(0.482)	(0.230)	(0.111)	(0.678)
Pre-Trial Test			0.452^{***}	0.456^{***}	0.219^{***}	0.229^{***}
			(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.686^{**}	−0.560	1.453^{*}	1.784^{**}	-2.244^{**}	−1.338
	(0.048)	(0.494)	(0.076)	(0.032)	(0.027)	(0.197)
Observations	2,301	2,301	2,301	2,301	1,898	1,898

Note: Full results for Table 4. Standard errors in parentheses, clustered at preschool level. *p-values* in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table A.6: Treatment effects on Math and Social Skills: Full Results with Covariates

	Math	Well-being Survey		
	(1) Math (Grade 3)	(2) General well-being	(3) Social skills	(4) Socio-em.distress
Treated × High school/less	0.181**	0.173*	−0.053	−0.325***
	(0.040)	(0.094)	(0.570)	(0.008)
Treated × College/more	−0.019	−0.071	0.056	−0.073
	(0.824)	(0.395)	(0.497)	(0.290)
College/more	0.624***	0.121	0.060	−0.231**
	(0.000)	(0.149)	(0.528)	(0.029)
Child age	0.143**	−0.079	0.033	−0.023
	(0.026)	(0.105)	(0.535)	(0.681)
Born in 2007	0.244	0.016	0.104	0.212
	(0.348)	(0.964)	(0.755)	(0.583)
Born in 2008	0.444	−0.202	0.059	0.252
	(0.109)	(0.581)	(0.853)	(0.519)
Born in 2009	0.685**	−0.180	0.127	0.167
	(0.016)	(0.627)	(0.693)	(0.678)
Born in 2010	0.611*	−0.400	0.109	0.093
	(0.076)	(0.297)	(0.755)	(0.827)
Male	0.043	−0.370***	0.237***	−0.064
	(0.312)	(0.000)	(0.000)	(0.254)
Birth weight (kg)	0.089	−0.059	0.052	−0.015
	(0.113)	(0.398)	(0.406)	(0.815)
Gestation (wks)	0.014	0.026	0.034*	−0.013
	(0.473)	(0.258)	(0.066)	(0.476)
Apgar score	−0.031	0.054	0.060	−0.069
	(0.451)	(0.167)	(0.228)	(0.169)
Number of Siblings	0.027	0.025	0.058*	−0.003
	(0.427)	(0.425)	(0.063)	(0.911)
Mother weight (kg)	−0.001	0.002	−0.000	−0.000
	(0.384)	(0.191)	(0.759)	(0.798)
Mother age	0.002	−0.015**	0.002	−0.005
	(0.768)	(0.041)	(0.756)	(0.547)
Mother employed	0.141*	0.219**	0.128	−0.041
	(0.070)	(0.015)	(0.123)	(0.567)
Father age	−0.003	0.002	0.004	0.003
	(0.585)	(0.710)	(0.468)	(0.616)
Father employed	0.034	−0.097	0.023	−0.070
	(0.690)	(0.228)	(0.843)	(0.504)
Household inc(1,000 USD)	0.001	0.001*	0.001**	−0.001**
	(0.111)	(0.092)	(0.033)	(0.020)
Constant	−2.700**	−0.310	−3.196***	1.826*
	(0.010)	(0.785)	(0.001)	(0.057)
Observations	1,635	1,339	1,339	1,339

Note: Full results for Table 5. P-values in parentheses, based on standard errors clustered at preschool level. P-values in parentheses. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Table A.7: Treatment-Control Differences in Parental Investments: Full Results with Covariates

	(1) Reading Inv.	(2) Non-cog Inv.
Treated × High school/less	0.183*	0.195*
	(0.089)	(0.057)
Treated × College/more	0.100	−0.145*
	(0.232)	(0.072)
College/more	0.250***	0.037
	(0.006)	(0.716)
Child age	−0.110**	−0.062
	(0.042)	(0.262)
Born in 2007	−0.998***	−1.013***
	(0.000)	(0.000)
Born in 2008	−0.782***	−0.833***
	(0.000)	(0.000)
Born in 2009	−0.692***	−0.918***
	(0.002)	(0.000)
Born in 2010	−0.730***	−1.018***
	(0.008)	(0.000)
Male	0.019	0.090
	(0.732)	(0.111)
Birth weight (kg)	−0.176***	−0.145**
	(0.004)	(0.022)
Gestation (wks)	0.027	−0.008
	(0.234)	(0.669)
Apgar score	0.008	−0.024
	(0.864)	(0.581)
Number of Siblings	−0.072**	−0.019
	(0.047)	(0.669)
Mother weight (kg)	−0.001	0.003
	(0.725)	(0.103)
Mother age	0.011	−0.016**
	(0.184)	(0.030)
Mother employed	0.155	−0.198**
	(0.125)	(0.013)
Father age	0.006	0.009
	(0.292)	(0.227)
Father employed	0.023	0.191
	(0.825)	(0.133)
Household inc(1,000 USD)	−0.001**	−0.001
	(0.023)	(0.142)
Constant	0.328	2.542***
	(0.774)	(0.006)
Observations	1,338	1,338

Note: Full results for Table 6. P-values in parentheses, based on standard errors clustered at preschool level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table A.8: Treatment-Control Differences in Parental Reading Investments: Split Factor

	(1) Inv: Time reading	(2) Inv: Enjoy reading
Treated × High school/less	0.157	0.068
	(0.131)	(0.123)
Treated × College/more	0.017	0.087
	(0.069)	(0.074)
College/more	0.115	0.255^{**}
	(0.097)	(0.110)
Child age	-0.140^{***}	−0.029
	(0.047)	(0.059)
Born in 2007	-1.429^{***}	−0.070
	(0.221)	(0.121)
Born in 2008	-1.157^{***}	−0.044
	(0.223)	(0.134)
Born in 2009	-1.044^{***}	−0.086
	(0.237)	(0.162)
Born in 2010	-1.348^{***}	0.174
	(0.280)	(0.214)
Male	0.018	0.039
	(0.053)	(0.057)
Birth weight (kg)	-0.161^{**}	−0.097
	(0.069)	(0.063)
Gestation (wks)	0.013	0.021
	(0.025)	(0.020)
Apgar score	−0.044	0.056
	(0.038)	(0.057)
Number of Siblings	−0.034	-0.079^{**}
	(0.042)	(0.037)
Mother weight (kg)	0.001	−0.001
	(0.002)	(0.002)
Mother age	−0.001	0.015^{**}
	(0.009)	(0.007)
Mother employed	0.053	0.106
	(0.101)	(0.092)
Father age	0.011	−0.001
	(0.007)	(0.006)
Father employed	0.039	0.032
	(0.132)	(0.106)
Household inc(1,000 USD)	-0.001^{**}	−0.000
	(0.000)	(0.000)
Constant	2.300^{**}	−1.468
	(1.142)	(1.093)
Mean outcome	−0.009	0.024
Observations	1,338	1,338

Note: Regressing split version of the parental reading-investment factor on a treatment indicator interacted with parental education, similar to Table 6. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.9: Heterogeneous Treatment Effects on Language Skills: Full Results with Covariates

	(1)	(2)
Treated × High school/less	0.232**	
	(0.021)	
Treated × College/more	−0.032	
	(0.681)	
Treated × School quality low		0.266
		(0.104)
Treated × School quality high		−0.015
		(0.850)
College/more	0.480***	
	(0.000)	
School quality high		0.336**
		(0.040)
Child age	0.041	0.013
	(0.485)	(0.821)
Born in 2007	0.064	0.123
	(0.778)	(0.607)
Born in 2008	0.191	0.215
	(0.407)	(0.367)
Born in 2009	0.241	0.272
	(0.359)	(0.314)
Born in 2010	0.289	0.290
	(0.348)	(0.348)
Male	−0.175***	−0.163***
	(0.000)	(0.000)
Birth weight (kg)	0.109**	0.103**
	(0.013)	(0.021)
Gestation (wks)	−0.005	−0.002
	(0.786)	(0.919)
Apgar score	−0.031	−0.034
	(0.223)	(0.135)
Number of Siblings	0.031	0.026
	(0.244)	(0.342)
Mother weight (kg)	−0.002	−0.002
	(0.259)	(0.161)
Mother age	0.018***	0.023***
	(0.007)	(0.001)
Mother employed	0.098	0.138*
	(0.161)	(0.051)
Father age	−0.001	0.001
	(0.879)	(0.853)
Father employed	0.075	0.095
	(0.321)	(0.238)
Household inc(1,000 USD)	−0.000	0.000
	(0.814)	(0.431)
Pre-Trial Test	0.232***	0.256***
	(0.000)	(0.000)
Constant	−1.238	−1.414
	(0.251)	(0.195)
Observations	1,841	1,841

Note: Full results for Table 7. P-values in parentheses, based on standard errors clustered at preschool level. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

Table A.10: Heterogeneous Treatment Effects by School Quality (Test Scores): Full Results with Covariates

	(1) Language Test (Gr.2)	(2) Reading Inv.	(3) Non-cog Inv.
Treated × High school/less × School quality low	0.332***	0.497***	0.282
	(0.009)	(0.009)	(0.116)
Treated × High school/less × School quality high	0.176	0.043	0.061
	(0.161)	(0.759)	(0.644)
Treated × College/more × School quality low	0.190	0.019	0.028
	(0.457)	(0.897)	(0.872)
Treated × College/more × School quality high	−0.085	0.075	−0.234**
	(0.234)	(0.485)	(0.023)
High school/less × School quality low	0.000	0.000	0.000
	(.)	(.)	(.)
High school/less × School quality high	0.146	0.277*	0.270
	(0.298)	(0.066)	(0.148)
College/more × School quality low	0.268	0.414***	0.187
	(0.100)	(0.006)	(0.313)
College/more × School quality high	0.649***	0.462***	0.180
	(0.000)	(0.003)	(0.276)
Child age	0.036	−0.075	−0.052
	(0.537)	(0.197)	(0.422)
Born in 2007	0.057	−1.006***	−0.940***
	(0.799)	(0.000)	(0.000)
Born in 2008	0.178	−0.784***	−0.805***
	(0.430)	(0.004)	(0.001)
Born in 2009	0.233	−0.634**	−0.843***
	(0.368)	(0.019)	(0.001)
Born in 2010	0.273	−0.654**	−0.960***
	(0.365)	(0.044)	(0.002)
Male	−0.171***	−0.002	0.008
	(0.000)	(0.980)	(0.880)
Birth weight (kg)	0.107**	−0.164**	−0.101
	(0.015)	(0.015)	(0.148)
Gestation (wks)	−0.006	0.018	−0.008
	(0.751)	(0.468)	(0.683)
Apgar score	−0.033	−0.000	−0.046
	(0.182)	(0.996)	(0.269)
Number of Siblings	0.032	−0.064	−0.016
	(0.225)	(0.105)	(0.735)
Mother weight (kg)	−0.001	−0.001	0.002
	(0.338)	(0.793)	(0.319)
Mother age	0.017**	0.007	−0.018**
	(0.014)	(0.495)	(0.019)
Mother employed	0.099	0.116	−0.133
	(0.147)	(0.284)	(0.139)
Father age	−0.000	0.005	0.012
	(0.993)	(0.472)	(0.137)
Father employed	0.073	−0.043	0.171
	(0.332)	(0.702)	(0.200)
Household inc(1,000 USD)	−0.000	−0.001*	−0.001*
	(0.437)	(0.075)	(0.093)
Pre-Trial Test	0.234***		
	(0.000)		
Constant	−1.218	0.438	2.419**
	(0.258)	(0.728)	(0.016)
Observations	1,841	1,129	1,129

Note: Full results for Table 8. P-values in parentheses, based on standard errors clustered at preschool level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

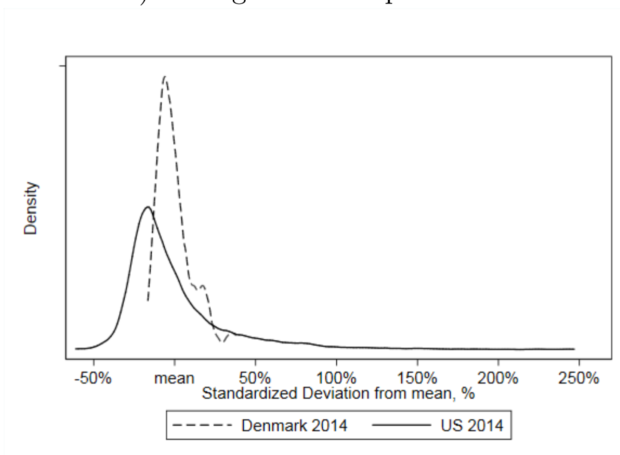
Table A.11: p-values of one-sided tests of hypothesis comparing treatment effects

	Treatment Effects on Language Test Scores $\frac{d\theta_2}{d\theta_1}$		
	$\theta_P^{low}, \underline{G}_2^{high}$	$\theta_P^{high}, \underline{G}_2^{low}$	$\theta_P^{high}, \underline{G}_2^{high}$
θ_P^{low}, $\underline{G}_2^{low}$	.189	.268	.003
θ_P^{low}, $\underline{G}_2^{high}$		.520	.011
θ_P^{high}, $\underline{G}_2^{low}$			.150
	Treatment Effects on Reading Investments $\frac{\partial x_2}{\partial \theta_1}$		
	$\theta_P^{low}, \underline{G}_2^{high}$	$\theta_P^{high}, \underline{G}_2^{low}$	$\theta_P^{high}, \underline{G}_2^{high}$
θ_P^{low}, $\underline{G}_2^{low}$	.030	.078	.038
θ_P^{low}, $\underline{G}_2^{high}$		.587	.630
θ_P^{high}, $\underline{G}_2^{low}$			.506
	Treatment Effects on Non-cognitive Investments $\frac{\partial x_2}{\partial \theta_1}$		
	$\theta_P^{low}, \underline{G}_2^{high}$	$\theta_P^{high}, \underline{G}_2^{low}$	$\theta_P^{high}, \underline{G}_2^{high}$
θ_P^{low}, $\underline{G}_2^{low}$	.201	.228	.004
θ_P^{low}, $\underline{G}_2^{high}$		.489	.017
θ_P^{high}, $\underline{G}_2^{low}$			.045

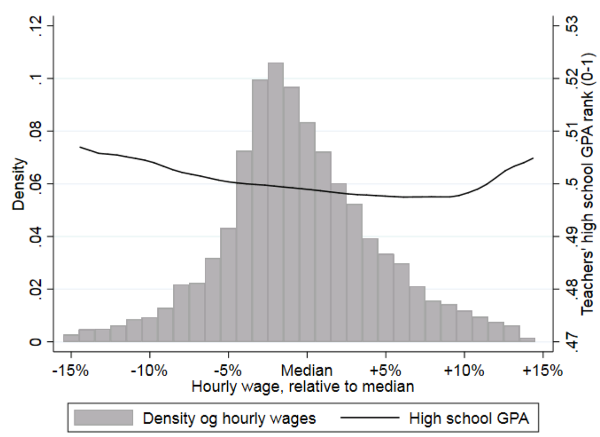
Note: These p-values compare treatment effects in Table 8 to each other. The tests performed are of the direction *row→column*. For example, the p-value of .077 in the top left cell corresponds to a test of H_0 : Treatment effect in $\theta_P^{low}, \underline{G}_2^{low}$ > Treatment effect in $\theta_P^{low}, \underline{G}_2^{high}$.

Figure A.1: Average Public School Expenditures in Denmark and the U.S.

a) Average school expenditure



b) Teachers' hourly wages



Note: Figure a) shows average school expenditures per student in public schools in 2014 relative the to the country average. Source: Denmark: www.statistikbanken.dk (Statistics Denmark); U.S.: Annual Survey of School System Finances; https://catalog.data.gov/dataset/annual-survey-of-school-system-finances. Figure b) shows the distribution of teachers' hourly wage rates in 2014 as a percentage deviation from the median wage rate. The figure also presents the association between teachers' rank of high school GPA and hourly wages (note that the y-axis only span from 0.47–0.53; $corr(wage, testscore) = -0.03$, with $p = 0.73$ for H_0 that $corr = 0$ and H_A that $corr \neq 0$). Hourly wage rates are adjusted for years of experience to remove variation stemming from the wage-progression at different levels of experience set by collective bargaining. This adjustment involves some measurement error, as it uses *years since graduation* and not *years of employment as a teacher in a Danish municipality*. Also, the hourly wage rates are not adjusted for the roughly 5% wage differences across regions (a PPP adjustment).

Table A.12: Heterogeneous Treatment Effects — Comparing School Quality Measures

	(1) School Quality Based on Average Test Scores (as Table 8)	(2) School Quality Based on Average Teacher Characteristics
Treated × High school/less × School quality low	0.332***	0.403**
	(0.124)	(0.166)
Treated × High school/less × School quality high	0.176	0.126
	(0.124)	(0.120)
Treated × College/more × School quality low	0.190	0.399
	(0.254)	(0.248)
Treated × College/more × School quality high	−0.085	−0.164**
	(0.071)	(0.065)
High school/less × School quality low	0.000	0.000
	(.)	(.)
High school/less × School quality high	0.146	0.173
	(0.140)	(0.145)
College/more × School quality low	0.268	0.290*
	(0.161)	(0.149)
College/more × School quality high	0.649***	0.671***
	(0.131)	(0.126)
Covariates	X	X
Observations	1,841	1,692

Note: Showing results from regressions of language test scores in grade 2 on the interaction of treatment status × parental education × school quality, where school quality is measured as in the main text (column 1), and with the alternative measure using predicted test scores based on average teacher characteristics at each school (column 2). More details on the quality definition in Section C.4. Covariates included as in Table 8, including pre-intervention test scores.

Table A.13: Heterogeneous Treatment Effects — Comparing Realized to Default School

	(1) Realized School Quality Based on Test Scores (as Table 8)	(2) Default School Quality Based on Test Scores
Treated × High school/less × School quality low	0.332***	0.411***
	(0.124)	(0.110)
Treated × High school/less × School quality high	0.176	0.111
	(0.124)	(0.120)
Treated × College/more × School quality low	0.190	0.164
	(0.254)	(0.243)
Treated × College/more × School quality high	−0.085	−0.087
	(0.071)	(0.070)
High school/less × School quality low	0.000	0.000
	(.)	(.)
High school/less × School quality high	0.146	0.235*
	(0.140)	(0.141)
College/more × School quality low	0.268	0.366**
	(0.161)	(0.175)
College/more × School quality high	0.649***	0.691***
	(0.131)	(0.118)
Covariates	X	X
Observations	1,841	1,898

Note: Showing results from regressions of language test scores in grade 2 on the interaction of treatment status × parental education × school quality, where we contrast the realized school (column 1) as in the main text to the default school (column 2). Default school is the elementary school attended by the majority of children at each given preschool. School quality is measured as in the main text with average test scores. Covariates included as in Table 8, including pre-intervention test scores.

A.1 Results not conditioning on covariates

Table A.14: Baseline balancing, short-run, and longer-run treatment effects — Not conditioning on covariates

	Pre-Trial Test		Post-Trial Test		Language Test Grade 2	
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	By parent ed	Pooled	By parent ed	Pooled	By parent ed
Treated	0.073		0.345***		0.055	
	(0.088)		(0.084)		(0.089)	
Treated × High school/less		0.080		0.351***		0.253**
		(0.094)		(0.089)		(0.097)
Treated × College/more		0.089		0.357***		−0.033
		(0.094)		(0.092)		(0.074)
College/more		0.419***		0.300***		0.661***
		(0.068)		(0.066)		(0.073)
Constant	0.002	−0.271***	0.475***	0.279***	0.129*	−0.295***
	(0.062)	(0.065)	(0.060)	(0.062)	(0.074)	(0.077)
Parental Education	-	-	-	-	-	-
Pre-test	-	-	-	-	-	-
Covariates	-	-	-	-	-	-
Observations	2,301	2,301	2,301	2,301	1,898	1,898

Note: Similar to Table 4, but not including covariates. Regression estimates of the treatment–control differences (β_t) in test scores y_{it} from $y_{it} = \alpha + \beta_t T_i + \varepsilon_{it}$. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10$,$^{**}p < 0.05$,$^{***}p < 0.01$. Main analysis sample excluding children with no immigration background. No other covariates included.

Table A.15: Treatment effects — Not conditioning on Pre-Intervention Test Scores

	Post-Trial Test		Language Test Grade 2	
	(1)	(2)	(3)	(4)
	Pooled	By parent ed	Pooled	By parent ed
Treated	0.347***		0.078	
	(0.000)		(0.277)	
Treated × High school/less		0.340***		0.250***
		(0.000)		(0.008)
Treated × College/more		0.353***		−0.017
		(0.000)		(0.824)
College/more		0.223***		0.558***
		(0.001)		(0.000)
Child age	−0.097**	−0.100**	0.018	0.019
	(0.035)	(0.033)	(0.746)	(0.748)
Born in 2007	0.004	−0.036	0.212	0.136
	(0.981)	(0.833)	(0.377)	(0.585)
Born in 2008	−0.117	−0.157	0.247	0.176
	(0.455)	(0.342)	(0.300)	(0.487)
Born in 2009	−0.251	−0.295	0.310	0.229
	(0.171)	(0.122)	(0.239)	(0.408)
Born in 2010	−0.222	−0.273	0.414	0.333
	(0.292)	(0.206)	(0.172)	(0.293)
Male	0.002	0.000	−0.171***	−0.174***
	(0.965)	(0.993)	(0.000)	(0.000)
Birth weight (kg)	0.108**	0.118**	0.114**	0.135***
	(0.015)	(0.010)	(0.015)	(0.004)
Gestation (wks)	−0.017	−0.018	0.001	0.002
	(0.251)	(0.235)	(0.949)	(0.909)
Apgar score	0.015	0.012	−0.034	−0.031
	(0.743)	(0.778)	(0.167)	(0.246)
Number of Siblings	−0.091***	−0.095***	0.009	0.006
	(0.003)	(0.002)	(0.742)	(0.825)
Mother weight (kg)	−0.001	−0.001	−0.002	−0.002
	(0.649)	(0.436)	(0.286)	(0.161)
Mother education (yrs)	0.047***		0.072***	
	(0.000)		(0.000)	
Mother age	0.004	0.008	0.017***	0.020***
	(0.438)	(0.187)	(0.010)	(0.004)
Mother employed	0.044	0.078	0.055	0.095
	(0.446)	(0.185)	(0.443)	(0.189)
Father education (yrs)	0.025**		0.048***	
	(0.026)		(0.000)	
Father age	−0.004	−0.005	−0.000	−0.002
	(0.443)	(0.362)	(0.951)	(0.696)
Father employed	−0.033	−0.021	0.047	0.082
	(0.595)	(0.728)	(0.540)	(0.277)
Household inc(1,000 USD)	0.000	0.001*	−0.000	0.000
	(0.240)	(0.051)	(0.212)	(0.752)
Constant	0.692	1.529*	−2.612**	−1.483
	(0.442)	(0.090)	(0.012)	(0.166)
Observations	2,301	2,301	1,898	1,898

Note: Similar to Table 4, but not including pre-test scores. Regression estimates of the treatment–control differences (β_t) in test scores y_{it} from $y_{it} = \alpha + \beta_t T_i + \varepsilon_{it}$. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Main analysis sample excluding children with no immigration background. No other covariates included.

Table A.16: Treatment-Control Differences in Math and Social Skills — Not conditioning on covariates

	Math	Well-being Survey		
	(1) Math (Grade 3)	(2) General well-being	(3) Social skills	(4) Socio-em.distress
Treated × High school/less	0.159*	0.228**	−0.080	−0.308**
	(0.088)	(0.112)	(0.092)	(0.125)
Treated × College/more	−0.034	−0.049	0.039	−0.073
	(0.086)	(0.084)	(0.086)	(0.068)
College/more	0.681***	0.154*	0.169*	−0.306***
	(0.063)	(0.085)	(0.089)	(0.109)
Constant	−0.337***	−0.167*	−0.083	0.293***
	(0.063)	(0.096)	(0.073)	(0.106)
Covariates	-	-	-	-
Observations	1,635	1,339	1,339	1,339

Note: Similar to Table 5, but not including covariates. Regressing children's outcomes on treatment status and parental educational attainment. Main analysis sample excluding children with an immigration background. A higher Socio-emotional distress score indicates a worse outcome. Standard errors (in parentheses) are clustered at the preschool level. $^{*}p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$.

Table A.17: Treatment–control Differences in Investments — Not conditioning on covariates

	(1) Reading Inv.	(2) Non-cog Inv.
Treated × High school/less	0.188	0.186*
	(0.115)	(0.107)
Treated × College/more	0.120	−0.122
	(0.093)	(0.079)
College/more	0.267***	−0.057
	(0.090)	(0.099)
Constant	−0.243***	0.024
	(0.085)	(0.081)
Covariates	-	-
Observations	1,338	1,338

Note: Similar to Table 6, but not including covariates. Standard errors (in parentheses) clustered at preschool level. $^{*}p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$.

Table A.18: Heterogeneous Treatment Effects by School Quality (Test Scores) — Not conditioning on covariates

	(1) Language Test (Gr.2)	(2) Reading Inv.	(3) Non-cog Inv.
Treated × High school/less × School quality low	0.316**	0.471**	0.215
	(0.128)	(0.211)	(0.172)
Treated × High school/less × School quality high	0.247**	0.070	0.062
	(0.119)	(0.146)	(0.136)
Treated × College/more × School quality low	0.151	−0.006	0.030
	(0.258)	(0.155)	(0.184)
Treated × College/more × School quality high	−0.066	0.089	−0.212**
	(0.066)	(0.116)	(0.100)
High school/less × School quality low	0.000	0.000	0.000
	(.)	(.)	(.)
High school/less × School quality high	0.112	0.264	0.245
	(0.127)	(0.166)	(0.181)
College/more × School quality low	0.407**	0.470***	0.136
	(0.177)	(0.138)	(0.173)
College/more × School quality high	0.809***	0.464***	0.058
	(0.117)	(0.156)	(0.163)
Covariates	-	-	-
Observations	1,841	1,129	1,129

Note: Similar to Table A.10, but not including covariates. Standard errors (in parentheses) clustered at preschool level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.2 Results on Full Sample, Including Children with Immigration Background

The following results replicate figures and tables from the main text, but they are run on the full sample of children (including those with an immigration background). The school quality measure is based on average test scores, as in the main text.

Table A.19: Number of Observations in different Treatment/Education/School Quality Groups — Children with Immigration Background Only

	(1) Full Sample	(2) Cond'l on Lang.Test Gr 2	(3) Cond'l on Lang Gr 2& Parent Survey
Control, Low ed, Low def. school qual	48	38	7
Control, Low ed, High def. school qual	11	10	3
Control, High ed, Low def. school qual	23	14	4
Control, High ed, High def. school qual	11	6	2
Treatment, Low ed, Low def. school qual	29	26	8
Treatment, Low ed, High def. school qual	17	15	7
Treatment, High ed, Low def. school qual	9	7	2
Treatment, High ed, High def. school qual	18	15	12
Total	166	131	45

Note: This table is equivalent to Table A.19, showing the added observations that are now present in the full sample that does not exclude children with an immigrant background.

Table A.20: Balancing of estimation sample by treatment — Including Children with Immigration Background

	(1) General Pop	(2) Control Avg.	(3) Treated Avg.	(4) Diff Treat-Control
Pre-Trial Test		0.039	-0.038	0.076
		(1.032)	(1.007)	(0.085)
Child Age at pre-trial test		4.047	4.098	-0.052
		(0.849)	(0.854)	(0.042)
Born in 2007	0.251	0.171	0.187	-0.017
	(0.434)	(0.376)	(0.390)	(0.045)
Born in 2008	0.255	0.335	0.332	0.003
	(0.436)	(0.472)	(0.471)	(0.024)
Born in 2009	0.246	0.330	0.301	0.029
	(0.431)	(0.470)	(0.459)	(0.022)
Born in 2010	0.247	0.154	0.171	-0.016
	(0.431)	(0.361)	(0.376)	(0.041)
Male	0.514	0.501	0.533	-0.033
	(0.500)	(0.500)	(0.499)	(0.025)
Birth weight (kg)	3.470	3.455	3.509	-0.053**
	(0.602)	(0.525)	(0.482)	(0.023)
Gestation (wks)	39.623	39.696	39.760	-0.064
	(1.938)	(1.583)	(1.563)	(0.076)
Apgar score	9.865	9.883	9.858	0.026
	(0.617)	(0.576)	(0.576)	(0.027)
Number of Siblings	1.450	1.340	1.444	-0.104*
	(0.954)	(0.810)	(0.941)	(0.059)
Mother weight (kg)	67.053	67.335	66.378	0.957
	(44.434)	(15.835)	(16.653)	(1.199)
Mother education (yrs)	13.762	14.036	13.994	0.042
	(2.635)	(2.487)	(2.714)	(0.315)
Mother age	38.941	39.747	39.823	-0.076
	(5.172)	(5.028)	(5.092)	(0.527)
Mother employed	0.729	0.831	0.797	0.034
	(0.445)	(0.375)	(0.402)	(0.030)
Father education (yrs)	13.558	13.777	13.833	-0.056
	(2.554)	(2.518)	(2.480)	(0.298)
Father age	41.590	42.025	42.580	-0.555
	(5.942)	(5.679)	(6.131)	(0.490)
Father employed	0.839	0.886	0.877	0.010
	(0.367)	(0.317)	(0.329)	(0.024)
Household inc(1,000 USD)	83.281	96.660	95.733	0.927
	(59.652)	(61.446)	(72.081)	(8.358)
School Quality	0.629	0.625	0.632	-0.007
	(0.252)	(0.247)	(0.259)	(0.055)
School Quality (Teacher Characteristics)	0.497	0.494	0.502	-0.009
	(0.318)	(0.317)	(0.320)	(0.073)
Observations	267,851	1,243	1,224	2,467

Note: The table shows descriptive statistics for all children in Denmark in the same birth cohorts as the paper's sample (column 1), the control group (column 2), and the treatment group (column 3). Average treatment-control differences are shown in column 4. Standard deviations of the variables are shown in parentheses for columns 1-3, standard errors clustered at institution level for column 4. The general population (column 1) consists of all children born in 2007-2010. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$. Note that the number of observations is only 2,186 for the 2 rows on school quality.

Table A.21: Treatment Effects Language Test Scores — Including Children with Immigrant Background

	Pre-Trial Test		Post-Trial Test		National Test Grade 2	
	(1) Pre-Trial Test	(2) Pre-Trial Test	(3) Post-Trial Test	(4) Post-Trial Test	(5) Language Test (Gr.2)	(6) Language Test (Gr.2)
Treated	0.071		0.308***		0.048	
	(0.337)		(0.000)		(0.506)	
Treated × High school/less		0.057		0.294***		0.193**
		(0.528)		(0.000)		(0.037)
Treated × College/more		0.089		0.318***		−0.035
		(0.322)		(0.000)		(0.641)
College/more		0.314***		0.064		0.467***
		(0.000)		(0.194)		(0.000)
Child age	−0.044	−0.044	−0.069*	−0.070*	0.033	0.038
	(0.337)	(0.335)	(0.071)	(0.071)	(0.544)	(0.498)
Born in 2007	0.051	−0.001	−0.142	−0.151	0.147	0.090
	(0.848)	(0.997)	(0.224)	(0.213)	(0.415)	(0.625)
Born in 2008	−0.201	−0.247	−0.120	−0.128	0.228	0.185
	(0.464)	(0.338)	(0.325)	(0.307)	(0.219)	(0.337)
Born in 2009	−0.198	−0.250	−0.252	−0.261	0.315	0.264
	(0.468)	(0.330)	(0.111)	(0.102)	(0.147)	(0.243)
Born in 2010	−0.046	−0.102	−0.291	−0.304	0.369	0.319
	(0.873)	(0.708)	(0.125)	(0.112)	(0.157)	(0.238)
Male	−0.016	−0.017	0.016	0.016	−0.186***	−0.186***
	(0.688)	(0.652)	(0.631)	(0.636)	(0.000)	(0.000)
Birth weight (kg)	0.080*	0.090*	0.057	0.060	0.105**	0.119**
	(0.089)	(0.065)	(0.133)	(0.124)	(0.027)	(0.011)
Gestation (wks)	0.008	0.007	−0.020	−0.021	−0.009	−0.008
	(0.622)	(0.680)	(0.141)	(0.136)	(0.619)	(0.659)
Apgar score	−0.022	−0.025	0.014	0.014	−0.032	−0.029
	(0.464)	(0.399)	(0.739)	(0.749)	(0.182)	(0.242)
Number of Siblings	−0.068**	−0.076***	−0.049**	−0.051**	0.030	0.023
	(0.012)	(0.005)	(0.046)	(0.035)	(0.220)	(0.344)
Mother weight (kg)	−0.000	−0.001	0.000	−0.000	−0.002	−0.003*
	(0.755)	(0.506)	(0.956)	(0.909)	(0.134)	(0.079)
Mother education (yrs)	0.054***		0.017**		0.057***	
	(0.000)		(0.043)		(0.000)	
Mother age	0.011**	0.014***	0.001	0.002	0.016**	0.017***
	(0.034)	(0.010)	(0.822)	(0.658)	(0.013)	(0.010)
Mother employed	−0.008	0.027	0.045	0.058	0.041	0.071
	(0.894)	(0.679)	(0.289)	(0.162)	(0.556)	(0.307)
Father education (yrs)	0.042***		0.009		0.040***	
	(0.000)		(0.341)		(0.000)	
Father age	−0.006	−0.008*	−0.001	−0.001	−0.002	−0.003
	(0.119)	(0.072)	(0.770)	(0.719)	(0.773)	(0.589)
Father employed	0.008	0.030	−0.045	−0.039	0.063	0.097
	(0.900)	(0.618)	(0.406)	(0.465)	(0.316)	(0.110)
Household inc(1,000 USD)	0.001	0.001**	0.000	0.000	−0.000	−0.000
	(0.180)	(0.017)	(0.311)	(0.142)	(0.162)	(0.911)
Immigrant Background	−0.292***	−0.328***	−0.076	−0.086	−0.027	−0.050
	(0.001)	(0.000)	(0.305)	(0.242)	(0.821)	(0.695)
Pre-Trial Test			0.451***	0.455***	0.218***	0.229***
			(0.000)	(0.000)	(0.000)	(0.000)
Constant	−1.356	−0.229	1.379*	1.685**	−1.852*	−0.936
	(0.101)	(0.775)	(0.079)	(0.035)	(0.055)	(0.345)
Observations	2,465	2,465	2,465	2,465	2,028	2,028

Note: Regression similar to Table 4 and Table A.5, but on sample including children with immigration background. P-values in parentheses, based on standard errors clustered at preschool level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table A.22: Treatment Effects, Testing Interaction with Immigrant Background

	Pre-Trial Test		Post-Trial Test		National Test Grade 2	
	(1) Pre-Trial Test	(2) Pre-Trial Test	(3) Post-Trial Test	(4) Post-Trial Test	(5) Language Test (Gr.2)	(6) Language Test (Gr.2)
Treated	0.084		0.309***		0.056	
	(0.270)		(0.000)		(0.468)	
Treated × Immigrant Background	−0.195		−0.012		−0.111	
	(0.244)		(0.937)		(0.634)	
Treated × High school/less		0.079		0.305***		0.223**
		(0.411)		(0.000)		(0.023)
Treated × College/more		0.091		0.311***		−0.040
		(0.314)		(0.000)		(0.603)
Treated × High school/less × Immigrant Background		−0.205		−0.098		−0.267
		(0.314)		(0.566)		(0.338)
Treated × College/more × Immigrant Background		−0.051		0.188		0.150
		(0.854)		(0.370)		(0.626)
Immigrant Background	−0.202**	−0.254**	−0.071	−0.027	0.028	0.081
	(0.022)	(0.046)	(0.422)	(0.817)	(0.850)	(0.640)
College/more × Immigrant Background		−0.017		−0.170		−0.217
		(0.937)		(0.286)		(0.328)
College/more		0.321***		0.078		0.487***
		(0.000)		(0.133)		(0.000)
Child age	−0.042	−0.044	−0.069*	−0.070*	0.034	0.040
	(0.361)	(0.348)	(0.071)	(0.069)	(0.535)	(0.484)
Born in 2007	0.063	0.003	−0.142	−0.151	0.152	0.092
	(0.815)	(0.991)	(0.216)	(0.200)	(0.405)	(0.627)
Born in 2008	−0.187	−0.241	−0.120	−0.127	0.235	0.191
	(0.493)	(0.346)	(0.318)	(0.292)	(0.209)	(0.331)
Born in 2009	−0.184	−0.244	−0.251	−0.262*	0.322	0.269
	(0.498)	(0.338)	(0.107)	(0.095)	(0.143)	(0.241)
Born in 2010	−0.028	−0.094	−0.290	−0.303	0.378	0.330
	(0.922)	(0.730)	(0.121)	(0.104)	(0.152)	(0.233)
Male	−0.016	−0.017	0.016	0.015	−0.186***	−0.187***
	(0.685)	(0.648)	(0.631)	(0.659)	(0.000)	(0.000)
Birth weight (kg)	0.080*	0.091*	0.057	0.060	0.105**	0.120**
	(0.090)	(0.065)	(0.134)	(0.126)	(0.026)	(0.011)
Gestation (wks)	0.008	0.007	−0.020	−0.020	−0.009	−0.008
	(0.625)	(0.678)	(0.141)	(0.140)	(0.617)	(0.663)
Apgar score	−0.021	−0.024	0.014	0.014	−0.031	−0.028
	(0.472)	(0.408)	(0.739)	(0.741)	(0.186)	(0.260)
Number of Siblings	−0.070***	−0.077***	−0.049**	−0.049**	0.030	0.024
	(0.010)	(0.005)	(0.045)	(0.045)	(0.227)	(0.312)
Mother weight (kg)	−0.000	−0.001	0.000	−0.000	−0.002	−0.003*
	(0.749)	(0.507)	(0.957)	(0.950)	(0.133)	(0.081)
Mother education (yrs)	0.055***		0.017**		0.058***	
	(0.000)		(0.043)		(0.000)	
Mother age	0.012**	0.015***	0.001	0.002	0.016**	0.017***
	(0.033)	(0.009)	(0.821)	(0.665)	(0.012)	(0.010)
Mother employed	−0.006	0.030	0.045	0.058	0.042	0.072
	(0.929)	(0.647)	(0.289)	(0.172)	(0.542)	(0.303)
Father education (yrs)	0.042***		0.009		0.040***	
	(0.000)		(0.343)		(0.000)	
Father age	−0.007	−0.008*	−0.001	−0.002	−0.002	−0.003
	(0.105)	(0.067)	(0.768)	(0.713)	(0.757)	(0.571)
Father employed	0.010	0.030	−0.045	−0.040	0.066	0.098
	(0.869)	(0.602)	(0.412)	(0.466)	(0.307)	(0.118)
Household inc(1,000 USD)	0.001	0.001**	0.000	0.000	−0.000	−0.000
	(0.181)	(0.017)	(0.311)	(0.145)	(0.159)	(0.914)
Pre-Trial Test			0.451***	0.455***	0.218***	0.229***
			(0.000)	(0.000)	(0.000)	(0.000)
Constant	−1.392*	−0.258	1.377*	1.669**	−1.874*	−0.986
	(0.091)	(0.747)	(0.080)	(0.036)	(0.055)	(0.325)
Observations	2,465	2,465	2,465	2,465	2,028	2,028

Note: Testing whether treatment effects are different for children with an immigration background. See Table A.21. P-values in parentheses, based on standard errors clustered at preschool level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

B Model of Skill Formation with Parental and Public Investments

B.1 The General Model

This appendix discusses a model of skill formation with public and private investments, as outlined in Section 5.1. The technology of skill formation we use begins with a standard production function for univariate end-of-period skills θ_t that includes self-productivity from θ_{t-1} and investments I_t:

$$\theta_t = j(\theta_{t-1}, I_t) \tag{B.1}$$

Investment is itself a function of parental direct investments P_t and skill investments via the institutional setting, or public G_t, where parents shape $P_t = p(x_t, \theta_P)$ through direct time investments x_t, of which the efficacy depends on their own skills θ_P. Public investments are a function of the neighborhood, which parents buy into via their wages, which are a function of their skills. The public may also decide to invest exogenously with $\underline{G}_t$ (this is where the intervention will happen). Public investments are thus a function $G_t = g(\theta_P, \underline{G}_t)$:

$$I_t = m\left(p(x_t, \theta_P), g(\theta_P, \underline{G}_t)\right). \tag{B.2}$$

Parents derive utility from their child's future skills, as well as from their own contemporaneous and future consumption and leisure. Parents spend their available time of 1 on child investments x_t, work in the labor market h_t, and leisure l_t:

$$1 = h_t + x_t + l_t \quad \text{for} \quad t \in \{1, 2\} \tag{B.3}$$

Without borrowing, parents' budget constraints each period are given by skill-specific wage rate $w\left(\theta_P\right)$:

$$c_t = h_t w\left(\theta_P\right) \tag{B.4}$$

If we consider a two-period model, parental **utility** is the following function of parental consumption and children's future skills:

$$U(c_1, c_2, l_1, l_2, \theta_2) = u_1(c_1, l_1) + \beta u_2(c_2, l_2) + \beta^2 V\left(\theta_2\left(\theta_1\left(\theta_0, I_1(x_1, \theta_P, \underline{G}_1)\right), I_2(x_2, \theta_P, \underline{G}_2)\right)\right) \tag{B.5}$$

where β is the discount factor. Parents maximize this utility, subject to the technology described in Eqs. (B.1) and (B.2) and time and budget constraints in Eqs. (B.3) and (B.4). The Lagrangian for this problem (ignoring the non-negativity constraints on time use) is

$$
\begin{aligned}
\mathcal{L}_{c_1,c_2,h_1,h_2,x_1,x_2,l_1,l_2} = & \; u_1(c_1,l_1) + \beta u_2(c_2,l_2) + \beta^2 V\left(\theta_2\left(\theta_1\left(\theta_0, I_1(x_1,\theta_P,\underline{G}_1)\right), I_2(x_2,\theta_P,\underline{G}_2)\right)\right) \\
& + \lambda_1\left(h_1 w\left(\theta_P\right) - c_1\right) \\
& + \lambda_2\left(h_2 w\left(\theta_P\right) - c_2\right) \\
& + \lambda_3\left(1 - h_1 - x_1 - l_1\right) \\
& + \lambda_4\left(1 - h_2 - x_2 - l_1\right)
\end{aligned}
$$

The straightforward first-order-conditions can be combined to yield the following equilibrium conditions in period 2:

$$
\beta \frac{\partial V}{\partial \theta_2} \frac{\partial \theta_2\left(\theta_1, x_2, \theta_P, \underline{G}_2\right)}{\partial x_2} = \frac{\partial u_2}{\partial l_2} \tag{B.6}
$$

$$
= w\left(\theta_P\right) \frac{\partial u_2}{\partial c_2}. \tag{B.7}
$$

In equilibrium, parents must be indifferent in allocating their time to direct investments in children (giving indirect utility through future child skills), additional leisure (giving direct utility), or consumption (converted to time-units via the multiplication with the wage rate).

B.1.1 Introducing the RCT in the model

An exogenous investment by the RCT will raise θ_1, child skills, without the parents adjusting their investments. This assumes that parents do not observe the increased skills right away, but only once they manifest at the end of the period. In the next period, parents can adjust their time allocation between investments and leisure as well as consumption to maintain the equilibrium condition in Eqs. (B.6) and (B.7).

Plain comparative statics can give an idea of how parents might react. A successful intervention will raise θ_1, and thereby θ_2, everything else equal (assuming only that skills are self-productive and not detrimental). This will lower the marginal utility from future child skills in $\partial V/\partial\theta_2$. Parents can lower the two right-hand sides by increasing consumption or leisure (consuming some of the benefits from exogenously higher future utility from higher child skills).

It is not certain ex ante, however, whether the full left-hand-side decreases, because increased θ_1 might increase or decrease the effectiveness of parental investments, $\partial\theta_2/\partial x_2$, depending on complementarities between θ_1 and the other terms in the production function. If the levels of parental investments, parental quality, and school quality are such that an increase in θ_1 *raises* the effectiveness of parental investments $\left(\frac{\partial^2\theta_2\left(\theta_1,x_2,\theta_P,\underline{G}_2\right)}{\partial x_2 \partial \theta_1} > 0\right)$, parents might not have to adjust, or even *increase* their investments if the productivity-effect is larger than the decreased marginal utility.

It could of course also be that θ_1 and x_2 are substitutes, such that $\frac{\partial^2\theta_2\left(\theta_1,x_2,\theta_P,\underline{G}_2\right)}{\partial x_2 \partial \theta_1} < 0$. In this case, parents should *reduce* investments by lowering x_2, or increase consumption or leisure.

B.1.2 Analyzing Long-run Treatment Effects of RCT

The long-run treatment effect of the RCT is given by the **total derivative** of long-run skills with respect to an exogenous change in skills from the first period, θ_1:

$$\frac{d\theta_2}{d\theta_1} = \underbrace{\frac{\partial \theta_2}{\partial \theta_1}}_{\substack{\text{Self-productivity} \\ \text{of increased } \theta_1}} + \underbrace{\frac{\partial \theta_2}{\partial x_2}}_{\substack{\text{Productivity} \\ \text{of parents'} \\ \text{time investments}}} \cdot \underbrace{\frac{\partial x_2^*}{\partial \theta_1}}_{\substack{\text{Re-optimization} \\ \text{of parents'} \\ \text{investments} \\ \text{after intervention}}} \tag{B.8}$$

(all evaluated at levels of x_2 before the change in θ_1).

From the treatment effects on parental investments, we know that only one group of parents adjusts their investments in reading skills: parents with low education and low school quality. For all other groups, changes in reading investments are not statistically significantly different from zero. The same is true for adjustments in non-cognitive investments, with the addition that the *reduction* in investments for highly educated parents with high school quality is statistically significant.

We now go through a few more detailed analyses of Eq. (B.8) for the four groups of children analyzed in the paper: parental education levels high and low ($\theta_P^{high}, \theta_P^{low}$), and school quality high and low ($\underline{G}_2^{high}, \underline{G}_2^{low}$).

$$\frac{\partial x_2^*}{\partial \theta_1} = + \text{ only for } \theta_P^{low}, \underline{G}_2^{low}, 0 \text{ for rest} \tag{B.9}$$

From the treatment effects on *long-run skills* in Table 8, we know that

$$\left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{low}} > \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} \text{ and } \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{high},\underline{G}_2^{low}} \text{ and } \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{high},\underline{G}_2^{high}} \tag{B.10}$$

$$\left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} = \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{high},\underline{G}_2^{low}} = \text{insig} \tag{B.11}$$

$$\left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} >^* \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{high},\underline{G}_2^{high}} \tag{B.12}$$

Evaluating the total derivative in Eq. (B.8) at different levels of parental quality and school quality, starting with Eq. (B.11):

$$\left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} = 0 \tag{B.13}$$

$$\Leftrightarrow \left.\frac{\partial\theta_2}{\partial\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} = -\left.\frac{\partial\theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} = -\left.\frac{\partial\theta_2}{\partial x_2}\right|_{\theta_P^{low},\underline{G}_2^{high}} \cdot 0 \tag{B.14}$$

$$\left.\frac{\partial\theta_2}{\partial\theta_1}\right|_{\theta_P^{low},\underline{G}_2^{high}} = 0 \tag{B.15}$$

where the equality to Eq. (B.14) stems from the fact that $\frac{\partial x_2^*}{\partial \theta_1}$ is not statistically significantly different from zero. Similarly,

$$\left.\frac{\partial \theta_2}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{low}} = 0 \tag{B.16}$$

These last two results imply that there is quite little direct self-productivity from θ_1 to θ_2. We take this with a grain of salt, because the point estimates for the total change in θ_2 were positive, just not statistically significant. We do take from this exercise, however, that without sustained parental investments, child skills do not self-produce to later periods in a major way.

From the contrast between treatment effects and changes in parental investments for the different education levels of parents within high-quality schools, we see that

$$\left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}} > \left.\frac{d\theta_2}{d\theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}} \tag{B.17}$$

$$\underbrace{\left.\frac{\partial \theta_2}{\partial \theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}}}_{=0} + \left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}} > \left.\frac{\partial \theta_2}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}} + \left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}} \tag{B.18}$$

$$\left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}} - \left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}} > \underbrace{\left.\frac{\partial \theta_2}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}}}_{\geq 0} \tag{B.19}$$

$$\Leftrightarrow$$

$$\left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}} > \left.\frac{\partial \theta_2}{\partial x_2}\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}} \tag{B.20}$$

$$\Leftrightarrow$$

$$\frac{\left.\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{low}, \underline{G}_2^{high}}}{\left.\frac{\partial x_2^*}{\partial \theta_1}\right|_{\theta_P^{high}, \underline{G}_2^{high}}} > \frac{\left.\frac{\partial \theta_2}{\partial x_2}\right|_{\theta_P^{high}, \underline{G}_2^{high}}}{\left.\frac{\partial \theta_2}{\partial x_2}\right|_{\theta_P^{low}, \underline{G}_2^{high}}} \tag{B.21}$$

From Eq. (B.21), we infer that the productivity of investments on skills ($\partial\theta_2/\partial x_2$) is smaller in $\theta_P^{high}, \underline{G}_2^{high}$ parents relative to $\theta_P^{low}, \underline{G}_2^{high}$ parents than the excess effect of the intervention on parental investments in $\theta_P^{low}, \underline{G}_2^{high}$ parents. This excess effect on the left-hand side ranges from zero (in reading) to positive (in non-cognitive investments). That is because reading investments are equally affected between $\theta_P^{low}, \underline{G}_2^{high}$ and $\theta_P^{high}, \underline{G}_2^{high}$ parents, and the changes in non-cognitive investments are *greater* in $\theta_P^{low}, \underline{G}_2^{high}$ than $\theta_P^{high}, \underline{G}_2^{high}$ parents (where they are actually significantly negative). This points to a greater productivity of investments in children of low-educated parents than highly educated parents, among children who face good quality schools in period 2. This finding is reasonably consistent with observed investment patterns.

C Data appendix

C.1 The Survey

In late April 2017, an invitation to participate in a survey was sent to all parents with children who had participated in the intervention. The invitation was sent via the personalized secure email *e-Boks*[29] to minimize non-response and ensure that all parents received the invitation. All residents in Denmark have such an email inbox in *e-Boks* and use this to receive (and send) official communications, such as from employers or public sector officials, on everything from children's medical visits, preschool and school enrolment, reception of public transfers, tax records, etc.

The letter is presented below.[30]

[29]See https://www.e-boks.com/danmark/en/what-is-e-boks/.

[30]In English:

Dear Parents to [Child]

We are a group of researchers who are studying the environments that help children flourish and provide them with the best possible beginning of their life.

Your child's daycare has been part of a project focussing on children's language development, and you were in this context asked to participate in a survey a couple of years ago.

The interplay between different activities in a child's day

We would like to request your assistance by filling our a similar questionnaire. The questionnaire ask questions relating to your everyday activities, habits, and how your view your child's everyday. We would like to ask you this to improve our understanding of how children's everyday activities in- and outside the home environment are linked.

The questionnaire can be found by following this link: LINK

It will at most take 15 minutes to respond to the questionnaire, and you will – upon completion – participate in a lottery with the possibility of winning an iPad. Lottery-participation is not conditional on having participated in the old survey.

The study has been approved by the Danish Data Authorities (National IRB board) and *all information is confidential and will be anonymized.*

We hope you will participate in the survey and thereby provide an important contribution to the understanding of the early childhood of all children. If you have any questions, please let us know by writing to: startpaalivet econ.au.dk.

Sincerely,
Dorthe Bleses (Professor, TrygFonden's Centre for Child Research)
Rasmus Landersø (Rasmus Landersø, Senior Research, The Rockwool Foundation Research Unit)

Den 21. april 2017

Kære forældre til [*barns navn*]

Vi er en gruppe af forskere, der er i færd med at undersøge, hvordan børn får de bedste betingelser til at udvikle sig under opvæksten og den bedst mulige start på livet.

Jeres barns dagtilbud har tidligere været med i et projekt med fokus på børns sproglige udvikling, og i den forbindelse har I for ca. 2 [3, 4] år siden fået tilsendt et spørgeskema.

Samspillet mellem aktiviteter i børns hverdag
Vi vil nu bede jer om at hjælpe os igen ved at udfylde et lignende spørgeskema. Det handler om jeres hjem, vaner og opfattelse af jeres barns hverdag. Vi vil gerne spørge jer om dette for bedre at forstå samspillet mellem de aktiviteter, som børn laver i deres hverdag både ude og hjemme.

Spørgeskemaet findes på dette link: www.spørgeskema.dk.

Det tager kun ca. 15 minutter at besvare spørgeskemaet, og når I besvarer, deltager I samtidig i en lodtrækning om en iPad. I behøver ikke have besvaret det foregående spørgeskema for ca. 2 [3, 4] år siden for at besvare dette.

Undersøgelsen er godkendt af Datatilsynet, og *alle oplysninger behandles anonymt og fortroligt*.

Vi håber, at I kan hjælpe os, og derved give et vigtigt bidrag til at øge forståelsen af, hvordan samfundet bedst muligt kan hjælpe alle børn på vej i deres tidlige år. Hvis I har spørgsmål til projektet, kan I kontakte os på startpaalivet@econ.au.dk.

Venlig hilsen

Dorthe Bleses (Professor, TrygFondens Børneforskningscenter på Aarhus Universitet)
Rasmus Landersø (Seniorforsker, ROCKWOOL Fonden)

Following the letter, two reminders were sent to non-respondents, and later non-respondents were contacted by phone.

C.2 Survey Response

The survey-response rate was 60%. In the main results, we do include non-respondents (as we have information on all the remaining key variables from the initial post-trial test scores and subsequent full population register data).

Table A.4 in Appendix A presents estimation results from regressions of survey-response (0/1) on baseline characteristics and treatment status. Survey response is not random: As would be expected, respondents are parents of more highly skilled children (pre-test), and are employed.

Note that child skills are correlated with parental age and years of schooling, and family income (also shown after the data description in Table 3). Importantly, there are no significant differences in response rates by treatment status, interacted with parental education.

C.3 Data Construction

This section describes the data construction. The first step was to collect the data from the intervention (see Section 2.2) and transfer it to Statistics Denmark. Here, the data was anonymized (i.e. all social security numbers were changed to anonymized unique *pnr*-numbers) with a code facilitating the link between the intervention data and the register data using the anonymized pnr-numbers. A similar procedure was conducted once the survey data had been collected.

The register data encompasses the entire population of Denmark from 1980 to the present with parent identifiers and household identifiers, allowing us to link the children from the intervention to their parents. From the demographic register we also identify the children's country of origin, date of birth, and home addresses (all anonymized). We also link the children to the educational register. These data also include unique preschool and school identifiers (institution-numbers) allowing us to identify the institutions the children attend along with their peers at the same institutions.

C.4 Background Characteristics and Outcomes

The National Birth Register provides information on children's birth weight, gestation length, Apgar score, and mothers' weight at the time of pregnancy.

Using the parental identifiers, we also include information on parents' completed education from the educational register (referring to education in 2014), employment status from the labor market register (for the year 2017), and household income from the income register (based on tax authorities' information, for the year 2017). Parent and child ages are recorded for September 1st, 2017.

Child outcomes The pre- and post-trial tests are constructed from 50 items relating to sound discrimination, rhymes, word-segmentation, and letter identification. We standardize the tests (mean zero, standard deviation of 1) relative to the control group.

The grade 2 test scores are part of the compulsory national tests from grade 2 through 8 (with language tests in grades 2, 4, 6, 8). The tests focus on three underlying constructs: Reading comprehension, decoding, and language comprehension, and they take place near the end of the school year. The tests are performed on computers using an adaptive system in which questions are determined by the student's performance earlier in the test. The test is scored electronically without teacher input. Following Sievertsen et al. (2016) and Beuchert and Nandrup (2018), we standardize these three individual scores, take the simple average, and re-standardize them within year.

Parent outcomes We construct *parental investments* from a factor analysis with 26 items that describe parental activities and opinions. After extensive exploratory factor analysis, we perform a principal-component analysis with the number of factors limited to five, adding an oblique promax rotation with power 3. The estimates are reported in Table C.1. From these estimates, we predict five factor scores with Bartlett scores. The six statements/questions that load on the parental investment factor have six potential answers ranging from, for example, highly disagree to highly agree. We assign these answers values 1 to 6 in the factor analysis. Note that if we predict a parental investment factor score from a factor analysis that uses exclusively the parental investment items (instead of the full list of 26 as in Table C.1, the results are very similar. These two versions of a parental investment factor are correlated at .97.

Items that mainly load on parental reading investment factor

- How many times last week has your child been read to (or read with) at home?
- If your child can read, how often in the past week have you sat with your child while it read to you?
- How many times last week have you or your child read, not counting schoolwork?
- I think it is boring or difficult to read for my child.
- I enjoy reading for my child.
- I am often too busy or too tired to read to my child.

Items that mainly load on parental reading investment factor

- I do a lot to teach my child to focus, concentrate, and complete a task.
- When I play or read with my child, it is important to finish before we stop or start new things.
- During the last week, how often did you and your child do everyday activities together, such as cooking?
- How often did you talk with your child about what they have done in preschool/school in the last week?
- How many times during the last month have you talked to your child about how he/she is doing more generally?

Table C.1: Factor Loading Matrix of Parental Activities and Opinions

	Neg.Pub Eval	Parental Inv.	Growth Mindset	Home Capital	Non-cog Investments
How many times last week has your child been read to (or read with) at home?	0.075	0.709	-0.094	0.055	0.244
If your child can read, how often in the past week have you sat with your child ...	0.080	0.528	-0.108	-0.160	0.391
How many times last week have you or your child read, not counting schoolwork?	0.055	0.585	-0.084	0.190	0.159
I think it is boring or difficult to read for my child.	0.063	-0.620	-0.154	-0.020	0.155
I enjoy reading for my child.	-0.004	0.643	0.146	0.033	-0.068
I am often too busy or too tired to read to my child.	0.074	-0.696	0.037	0.044	-0.057
As a parent, I have a big influence on how my child is going to learn to read, ...	-0.003	0.092	0.566	-0.030	0.169
My child's ability to learn to read, count and calculate are intrinsic ...	-0.063	0.059	-0.568	-0.020	0.053
My child can always improve its ability to learn to read and count, no matter ...	0.055	-0.081	0.670	-0.030	0.168
After a certain time my child will no longer be able to improve its ability to ...	-0.010	0.110	-0.615	-0.057	0.043
I can affect my child's ability to focus on completing a task.	-0.005	0.026	0.727	-0.003	0.090
There is not much I can change if my child has a harder time concentrating.	0.048	-0.046	-0.672	-0.025	0.034
I do a lot to teach my child to focus, concentrate, and complete a task.	-0.056	0.086	0.166	-0.169	0.544
When I play or read with my child, it is important to finish before we stop ...	0.152	0.049	0.090	-0.195	0.375
During the last week, how often did you and your child do everyday activities ...	-0.077	0.039	-0.009	0.293	0.490
How often did you talk with your child about what they have done in preschool ...	-0.100	0.113	0.047	0.001	0.622
How many times during the last month have you talked to your child ...	-0.028	0.035	0.121	0.079	0.493
I think the amount my child is being read to in preschool(school) is not sufficient.	0.678	-0.066	0.008	0.069	0.011
I would like my child to receive more help to develop his/her language.	0.679	-0.084	-0.032	-0.009	0.120
How satisfied are you with the quantity of language support your child receives?	-0.787	-0.152	-0.038	0.010	0.293
How satisfied are you with the quality of language support your child receives?	-0.822	-0.149	-0.075	-0.029	0.276
One of the reasons I support my child's ability to focus, concentrate, ...	0.667	-0.098	-0.005	0.020	0.173
I would like my child to receive more help to develop his ability to concentrate	0.610	-0.109	-0.094	-0.017	0.186
How many books do you have in your home?	0.048	0.000	0.023	0.845	-0.078
How many children's books do you have in your home?	0.025	0.117	-0.025	0.757	-0.013
In the last week, how many times did you read books, newspapers, e-books, ...?	-0.034	-0.050	0.039	0.612	0.221

Note: Factor loadings after PCA on all 26 items listed here, limited to 5 factors, with oblique promax rotation (power 3). $N = 1,336$. "Neg.Pub.Eval." stands for a negative evaluation of the public investments by parents. "Parental Inv." is the parental direct time investment factor used in the main analyses. "Growth Mindset" relates to how parents view their child's potential to change, and their own potential to influence their child's growth in both the cognitive and non-cognitive domains. "Home Capital" relates to the capital present in the home that could foster reading and language. "Noncog. Important" describes how important it is for parents to foster their child's socio-emotional skills, in addition to reading and language.

Finally, we construct the variable on *hours worked* from survey responses to the following questions *At what time do you usually go to work?* and *At what time do you usually leave work?*.

School quality Our main measure of school quality is the rank of a school in terms of average test scores in Danish (grades 2,4,6, and 8) and Math (grades 3 and 6), in the years 2010-2016. These are preceding the years in which any RCT participants would be in elementary school.

We also generate an alternative measure of school quality that is based on the average characteristics of the teachers employed in each school in Denmark. We use a unique link developed by Statistics Denmark between all teachers (their pnr-numbers) and schools (institution-numbers) using employment records from the employer-employee match data to identify the full set of teachers employed at each school by January 1st from 2010-2016.

We link this data with the educational register, labor market register, and GPA from high school and teachers' college (UDG) to construct variables with each teacher's years of experience, tenure at a given school, unemployment spells and periods with sick leave, and GPA from high school and teachers' college.

The institution identifiers allow us to merge the aforementioned data with children's national test scores (see earlier paragraph). We obtain the predicted test scores from teacher characteristics by regressing the children's test scores on the school-by-year average teacher information. Finally, we rank schools from lowest to highest (0-1) by their predicted test score level.

Danish Well-being Survey The measures on well-being and socio-emotional skills come from students' answers on a large, national survey, "The Danish Wellbeing Survey." This is a yearly survey that is mandatory for public schools to administer since 2015. It is typically administered electronically during one class session. The announced purpose of the survey is to improve the well-being of all students at the school. Students are told that their individual responses will not be shown to their parents, teacher, or other persons at the school, that they should respond honestly, and that they could have the questions read aloud if they had reading problems or be helped in other ways. The questionnaires are linked to the students' national identification number, unless parents asked for their children to be anonymous—an option that exists since 2018.

Items that mainly load on General Well-Being

- Are you happy with your school?
- Are you happy with your class?
- Are you happy with your teachers?
- Can you concentrate in class?
- Are the teachers good at helping you in school?
- Are the classes boring?

Table C.2: Factor Structure in Well-Being Survey Grades 0-3

	General well-being	Social skills	Socio-em. Distress
Are you happy with your school?	0.579	0.166	0.030
Are you happy with your classroom?	0.372	0.315	0.068
Do you feel alone at school?	-0.108	0.262	0.537
Do you like the breaks at your school?	0.022	0.415	0.113
Are you happy with your teachers?	0.738	-0.027	-0.076
Does your stomach hurt when you're in school?	0.048	-0.104	0.699
Does your head hurt when you're in school?	0.141	-0.189	0.684
Are you good at solving your problems?	-0.032	0.639	-0.017
Can you concentrate in class?	0.260	0.244	0.180
Are you and your classmates good at helping each other?	0.224	0.537	-0.090
Do you think the other children in your class like you?	0.009	0.616	0.107
Are the teachers good at helping you in school?	0.631	0.068	-0.006
Is there anyone who is teasing you so you get sad?	-0.131	0.237	0.592
Are you afraid the other children will laugh at you?	-0.174	0.204	0.563
Do you get to say what you are doing in class?	0.043	0.471	-0.193
Are the classes boring?	0.651	-0.148	0.170
Do you learn something interesting in school?	0.662	0.091	-0.106
Is it difficult to hear what the teacher says in class?	0.165	-0.068	0.494

- Do you learn something interesting in school?

Items that mainly load on Social Skills

- Do you like the breaks at your school?
- Are you good at solving your problems?
- Are you and your classmates good at helping each other?
- Do you think the other children in your class like you?
- Do you get to say what you are doing in class?

Items that mainly load on Socio-emotional Distress

- Do you feel alone at school?
- Does your stomach hurt when you're in school?
- Does your head hurt when you're in school?
- Is there anyone who is teasing you so you get sad?
- Are you afraid the other children will laugh at you?
- Is it difficult to hear what the teacher says in class?

Sample Sizes For an overview of the availability of test scores and other outcomes, as well as parental investments, see Table C.3, with further breakdowns by treatment status presented in Table C.4.

Table C.3: Sample Sizes for Different Outcomes

	Individual Samples		
	mean	sd	count
Pre-Trial Test	0.039	1.009	2,301
Post-Trial Test	0.647	0.957	2,301
Language Test Gr.2	0.156	0.993	1,898
Math (Grade 3)	0.105	1.005	1,635
General well-being	-0.041	0.999	1,339
Social skills	0.021	0.977	1,339
Socio-em.distress	0.019	0.980	1,339
Reading Investment	0.013	0.990	1,338
Non-cognitive Investment	-0.012	0.998	1,338
	Conditional on Language Test Grade 2		
	mean	sd	count
Math (Grade 3)	0.112	1.002	1,602
General well-being	-0.034	0.982	1,321
Social skills	0.024	0.978	1,321
Socio-em.distress	0.016	0.967	1,321
Reading Investment	-0.012	0.984	1,103
Non-cognitive Investment	-0.024	1.002	1,103
	Conditional on Parent Survey		
	mean	sd	count
Math (Grade 3)	0.193	1.016	941
General well-being	-0.028	0.985	778
Social skills	0.052	0.962	778
Socio-em.distress	-0.038	0.941	778

Note: For a breakdown of sample sizes by treatment status and parental education, see Table C.4.

Table C.4: Number of Observations in different Treatment/Education/School Quality Groups

	(1) Full Sample	(2) Cond'l on Lang.Test Gr 2	(3) Cond'l on Lang & Parent Survey
Control, Low ed, Low def. school qual	154	138	55
Control, Low ed, High def. school qual	246	205	114
Control, High ed, Low def. school qual	135	103	72
Control, High ed, High def. school qual	615	508	303
Treatment, Low ed, Low def. school qual	106	89	42
Treatment, Low ed, High def. school qual	330	289	153
Treatment, High ed, Low def. school qual	89	66	38
Treatment, High ed, High def. school qual	626	500	326
Total	2,301	1,898	1,103

Note: Showing sample sizes for the full analysis sample (column 1, excluding children with immigration background), restricting on availability of the long-run language outcome (column 2), and additionally also on availability of responses on the parent survey (column 3).