EARLY CHILDCARE ATTENDANCE AND COGNITIVE SKILLS IN ADOLESCENCE

By

Ingvild Almås, Nina Drange, Costas Meghir and Henrik Daae Zachrisson

April 2026

COWLES FOUNDATION DISCUSSION PAPER NO. 2518

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Early Childcare Attendance and Cognitive skills in Adolescence*

Ingvild Almås[†], Nina Drange[‡], Costas Meghir[§] and Henrik Daae Zachrisson[¶]

April 2026

Abstract

This paper examines the impact of early childcare on academic achievement for children in grade 5 and grade 9, based on a 2003 policy expansion that created quasi-random variation in slot availability for children aged 1–2. Starting childcare one year earlier increases math scores by 9.7% of a standard deviation (SD) in grade 9. Children whose mothers do not hold a high school diploma benefit by a significant 28% of a SD at grade 9, reducing the math achievement gap from children of higher-educated mothers by about one third. We also present evidence of strong improvements for children of immigrants.

Keywords: Early Childhood Education and Care, Child Development, childcare, Test Scores
JEL codes: I21, I24, J13

*__Acknowledgement:__ Nina Drange was supported by ERC CoG Grant agreement No. 818425. Henrik Daae Zachrisson was supported by ERC CoG Grant agreements No. 818425 and 101045526, and The Research Council of Norway's Center for Research Excellence grant No. 331640. Costas Meghir thanks the Cowles foundation and the Institution for Social Policy Studies at Yale for financial support. The Norwegian registry data was used via the project SUBPU, which is approved by the Regional Committees for Medical and Health Research Ethics (ref. 2017/2205). The University of Oslo is responsible for the data handling in SUBPU and has conducted a Data Protection Impact Assessment (DPIA) in collaboration with the Norwegian Agency for Shared Services in Education and Research (Sikt; ref. 962088). The data access and management of SUBPU is financed by the Research Council of Norway, the European Research Council, and the Department of Psychology at the University of Oslo. This work was performed on the TSD (Tjenester for Sensitive Data) facilities, owned by the University of Oslo, operated and developed by the TSD service group at the University of Oslo, IT-Department (USIT). The usual disclaimer applies.

[†]Zurich University, ingvild.almas@econ.uzh.ch
[‡]Ragnar Frisch Centre for Economic Research, nina.drange@frisch.uio.no
[§]Yale University, IFS, CEPR and NBER, c.meghir@yale.edu
[¶]University of Oslo, h.d.zachrisson@isp.uio.no

1 Introduction

As women have been increasingly integrated in the labor market since the 1950s, childcare infrastructure has been growing in many western countries. There has also been increasing emphasis on providing childcare in an affordable manner, either by direct subsidies or by vouchers. Examples include the UK, the US, Canada, The Netherlands, Japan and Norway to mention but a few. While such infrastructure was mainly conceived as a way to increase career opportunities for mothers, the role that these can play in promoting child development, particularly in low socio-economic (SES) settings is at least as important. On the one hand, it is quite possible that childcare, particularly in the early years, is detrimental to child development (at least for some groups of children), but it is equally possible that well-designed childcare promotes child development. Indeed, there is some evidence for both arguments.

The emphasis on the early years is important, because it is a period characterized by rapid neurological changes and the growth of cognitive and non-cognitive skills (Nelson et al. 2024, Phillips & Shonkoff 2000). Later achievement and adult outcomes, including school performance and economic, demographic, and health outcomes, are based on foundations that are laid early in life (Heckman 2006, Cunha et al. 2010, Black et al. 2017). In particular, a safe, nurturing, and stimulating environment during the first 1,000 days from conception lays the foundation for subsequent development, while insufficiency in the early developmental context of a child can have lasting negative consequences (Daelmans et al. 2017). This period may therefore be seen as a window of opportunity: well-designed early interventions have the potential to generate large and lasting returns, and equalize opportunities (Currie 2001, Almond & Currie 2011), by helping to promote skill formation at a malleable stage.

The effects of the introduction of early childcare is likely to depend on the quality of the daycare offered, including the staff:child ratio, and also the quality of care that it replaces. We also observe that the empirical evidence on the effects of the introduction of early childcare is mixed, something that is likely to reflect different qualities of the care options, in daycare and outside, across different contexts (van Huizen & Plantenga 2018, Kuehnle & Oberfichtner 2020). Existing studies of universal- or subsidized large-scale programs measure short- and medium-term outcomes (measured either during preschool or the early school years). Some studies, most notably from Italy and Canada, report adverse effects on cognitive or behavioral outcomes (Fort et al. 2020, Baker et al. 2008, 2019). Others, including studies from Norway and Germany find neutral or positive effects (Dearing et al. 2015, Drange & Havnes 2019, Felfe & Lalive 2018, Zachrisson et al. 2024, Dearing et al. 2018), while a French study find negative effects on behavior but positive on cognitive outcomes (Berger et al. 2021). These contrasting findings underscore the importance of program content, quality, and socio-economic context in shaping developmental outcomes.

Most studies on the effects of childcare in the first years of life focus on outcomes in the preschool and early school years. Large-scale programs for infants and toddlers are relatively recent, and children who attended these programs are only now reaching the end of compulsory schooling, making the evaluation of longer-term effects possible. Current debates regarding the persistence or "fadeout" of effects from large-scale preschool programs for 3-5 year-olds (Burchinal et al. 2024, García 2024), as well as meta-analyses on fadeout in early childhood programs (Hart et al. 2024), further highlight the ongoing uncertainty about the longevity of early gains. To date, we lack evidence on the longer term impacts of large-scale universal programs for infants and toddlers across the SES spectrum. To our knowledge, no previous studies have tested effects for this age group on outcomes in adolescence with data allowing to differentiate effects for children from lower SES backgrounds. In this paper, we study long-term outcomes in terms of school performance at ages 10 to 14, using standardized testing for the multiple full cohorts of the Norwegian population. Norway is a resource rich and high trust society, where play based learning is in focus, and where the quality of the daycare offered could be considered as high, at least as measured by tightness of adults per child and the level of education required in the group of adults. The Norwegian setting at the start of the millennium can thus be seen as ideal for gathering new evidence on the possible effect of early daycare for children with different backgrounds and thus different outside care options.

In the aftermath of a political compromise, there was an extensive expansion of childcare slots for 1-2 year olds starting in 2003. Through large transfers to municipalities, the capacity doubled the coverage for 1-2 year-olds from about 40 to 80 percent over the period 2003-2007. Throughout this expansion, there were national guidelines to ensure quality and content. This included staff:child ratios of 1:3 (one of whom should have a tertiary degree in early childhood education), small group sizes (most commonly 9 children with 3 staff), and a national curriculum emphasizing a holistic, play-based, and child-centered approach to child development. Equipped with registry data, we can follow all children born in Norway and affected by the expansion throughout primary school and until they graduate from secondary school, thus adding new evidence on how early childcare affect cognitive skills throughout childhood and beyond.

We find that early start has a positive effect on math achievements at ages 10 to 14. The estimates suggest that starting childcare one year earlier increases math scores by 9.7 percent of a standard deviation. We find no corresponding overall effects on reading. Importantly, there is substantial heterogeneity by maternal education. For children whose mothers do not hold a high school diploma, early enrollment in childcare generates statistically significant mathematical test score improvements of 28 percent of a standard deviation. The magnitude of these effects suggests that starting childcare one year earlier reduces the math achievement gap between children of low- and higher-educated mothers by roughly one third. Reading scores for this group also improve, although these estimates are less precise. Finally, we also investigate the impacts of early childcare for immigrant children, and we find evidence of substantial improvements in math scores. Together, these results imply strong benefits for disadvantaged children and no harm for children from higher SES backgrounds.

The paper proceeds as follows. In Section 2 we review the most relevant previous evaluations of childcare reforms for 1 and 2 year old children. In Section 3 we explain the childcare system in Norway and describe the process of expansion. The administrative data sources we use are presented in Section 4 and our empirical strategy, which exploits the expansion, is discussed in Section 5. The results are discussed in Section 6 and Section 7 offers a summary and some concluding remarks.

2 Literature review

We frame our discussion of the previous literature primarily by focusing on evaluations of large-scale childcare programs for 1 and 2 year old children. Across these studies, differences in policy design, age of treatment, sample composition, age at which outcomes are measured, and identification strategy are central to understanding why estimated effects differ across contexts. We focus on results for cognitive development – that is the focus of this paper. However, we will also mention the importance of studying results for other developmental domains in the conclusion of this paper.

A number of highly influential studies are based on the introduction of Quebec's low-fee childcare program in 1997 (Baker et al. 2008, 2019, Kottelenberg & Lehrer 2017) which lead to a sharp increase in childcare coverage. The reform, as described by (Baker et al. 2008), initially set fees at $5 per day, and was phased in by age cohort, starting with the oldest children, and fully covered children aged 0–4 by 2000. The program was universal, hence children were eligible regardless of parents' employment status. However, prior to the 1997 reform, there had been extensive subsidies for low-income families. Therefore, the expansion of childcare coverage during the reform was most pronounced for higher-income families.

Baker, Gruber, and Milligan (2008) provided the initial large-scale evaluation of this reform. Using repeated cross-sections from the National Longitudinal Survey of Children and Youth (NLSCY) with several thousand children per wave, they used a difference-in-differences design comparing Quebec to other Canadian provinces before and after 1997. Outcomes were measured in early childhood (typically ages 2–4/5). They document that effects on cognitive outcomes were small and generally insignificant. Notably, they reported effects collapsed across age groups (entry age 0-4). The content of the childcare provided was in general of relatively low quality, however. The youngest children attended home-based care with regulated provider, who were required to have 24-45 hrs training and 6hrs annually professional development, and the observed quality was low (Japel et al. 2005).

Baker, Gruber, and Milligan (2019) extended this analysis to long-run outcomes using administrative data on crime and survey-based measures of well-being–again for children entering care across the age-span 0–4 years . Leveraging the same difference-in-differences framework and large population-level samples, they followed exposed cohorts into adolescence and early adulthood (up to approximately ages 15–20). They found positive treatment effects on PISA scores at age 15 of 20-30% of a standard deviation, but null effects on a Canadian national assessment on the same subject around the same age.

Using the same reform but a different empirical strategy, Kottelenberg and Lehrer (2017) reexamined early childhood outcomes in the NLSCY. Their sample was similar in size to those in the other Quebec studies, while the focus was on earlier entry into care (ages 0-3–hence more directly relevant to the age groups in our study), with outcomes measured in early childhood. Estimating heterogeneous effects, they showed that children from disadvantaged single-parent families, and especially those in the lower tail of the skill distribution, experience gains in cognitive measures, while children from more advantaged two-parent families experience negative effects.

Taken together, evidence from the Quebec reform suggests that a rapid, universal expansion affecting children from birth to age four may generate heterogeneous effects across sociodemographic groups, ranging from negative to positive. These results are only partially relevant for understanding long-term effects of childcare for the youngest children (only specifically addressed by Kottelenberg and Lehrer (2017)). Yet, they are important for recognizing that quality and content of the program relative to the counterfactual may be critical when interpreting childcare effects.

Berger, Panico, and Solaz (2021) used data from France to study short-term effects of attending publicly regulated crèches at age one. The system combines extensive public support with rationed access at the municipality level. Quality in the crèches is relatively high, with staff:child ratios of 1:5 for children not yet walking, and 1:8 for older children. Staff is required to have a secondary or university-level degree in a relevant discipline. Using data from a nationally representative birth cohort study (n of app. 12,000), they instrumented crèche attendance with interactions between quarter of birth and local crèche supply. Treatment was narrowly defined as center-based care at age one, and outcomes were measured at age two. Their instrumental variables estimates indicated positive effects on language development. Compared to Quebec, the French reform operated within a more tightly regulated system and isolates exposure at a specific age, highlighting the importance of both quality standards and the timing of treatment.

In Italy, Fort, Ichino, and Zanella (2020) evaluated municipal day care in Bologna, Italy. Demand for childcare exceeded supply and admission was administratively determined, allowing for a regression discontinuity design. Children entered childcare between ages 4 months and 2 years, in centers with staff:child ratios of 1:4 for children below age 1, and 1:6 for older children. There was a uniform curriculum and quality was regarded high, yet details are not described. Outcomes were measured between ages 8 and 14 in a sample of approximately 450 children (about 30% participation rate). The results showed that additional months of public day care reduced measured IQ for this sample. The regression discontinuity design identified local average treatment effects for children at the margin of admission, whose counterfactual care was typically with parents or grandparents. In this context, public provision appears to crowd out high-quality home inputs among advantaged families, generating negative long-run effects. Finally, Felfe and Lalive (2018) studied the expansion of childcare for children under age three in Germany during the 2000s. The reform expanded subsidized slots at different rates across districts in response to federal mandates. Using administrative and survey data covering several thousand children, they instrumented individual attendance with district-level variation in childcare availability to estimate marginal treatment effects. Children entered childcare between ages 0 and 2, while the quality and content of the programs are not described in detail. Results showed modest cognitive effects. Relative to Quebec and Bologna, the German expansion appeared to have drawn more disadvantaged children into care at the margin and occurred within a regulatory framework emphasizing quality standards.

There are some previous studies of early childcare in Norway. Using a large health survey (including about 100,000 children) which coincided with the roll-out of the reform we evaluate in this study with a similar IV strategy, Dearing et al. (2018) found large treatment gains (80% of a standard deviation) in language skills for 3 year old children from low-income families, with small or no effects on children from more affluent families. Drange & Havnes (2019) exploited a lottery allocating children to slots in public childcare. Children assigned by the lottery started publicly funded childcare at approximately age 1, while those not assigned started, on average, 6 months later. Those starting early scored 16% and 11% of a standard deviation better in reading and math screening tests in first grade (age 6). Moreover, for children of parents with fewer years of education, the effects were about 60% larger than the average effects, fully closing the differences observed between children from families with high- and low parental education in reading, while reducing it by 25% in math. Finally, Zachrisson et al. (2024) relied on the same sample and identification strategy as Dearing et al. (2018) found effects on national tests in 5th grade. The results were restricted to children of parents with few years of education, showing that early childcare reduced the gap in test scores with up to 50% .

Taken together, these studies underscore that the effects of large-scale childcare expansions depend critically on policy design, the age at which exposure occurs, the population induced into care, and the horizon over which outcomes are measured.

3 Institutional setting

The Norwegian childcare system offers universal coverage and is currently taken up by about 97 percent of children aged 3-5 years and about 90 percent of children aged 1-2 years.[1] The system was gradually built after the Second World War and has been expanded through several reforms from then and up until today.

In 1975 the Kindergarten Act ("Barnehageloven") established that day care is a public responsibility and defined the provision of such as a municipal responsibility. The system was designed as universal in scope in the sense that it was open to all

[1] Downloaded from *https : //www.ssb.no/utdanning/barnehager/statistikk/barnehager* on February 11, 2026.

children, regardless of income or social status, but it did not confer an individual legal right to a slot. Access depended on municipal capacity, and coverage expanded gradually over the following decades, particularly for children above the age of three. A decisive institutional shift occurred with a reform enacted in 2003, where an almost unified parliament agreed to a progression towards full provision of childcare for 1 and 2 year old children (the budgetary groundwork for which was laid out in the 2001-2002 National Budget and a White paper from the Government). This reform transformed the system from a universal service model to a rights-based entitlement by imposing a municipal obligation to guarantee a place for eligible children. Thereby, it consolidated universal coverage on a statutory basis.

It is the latter of these two reforms that we use to identify the effect of day care for children from the age of 1 years old. By the time of this reform, parents received up to one year of paid parental leave following the birth of a child. Consequently, very few children entered non-parental care before age 1. At that age, parents would typically either enroll their child in publicly subsidized childcare or receive cash-for-care benefits, which could be used to support staying home with the child or to finance informal childcare arrangements. The cash-for-care benefit amounted to approximately NOK 3,500 per month.[2] All parents with a child under 3 years of age who did not use subsidized public childcare were eligible, and parents were free to switch between the two options at their own discretion.

Throughout this period and to this day, the Norwegian public childcare system has been regulated to ensure high quality. One-third of the staff is required to hold a three-year college degree in early childhood education. In a typical classroom for children under three years old, a teacher and two assistants, usually untrained or holding a relevant secondary degree, are responsible for a group of 9 children, yielding a staff-

[2]USD 560 per month using the 2006 exchange rate (Ministry of Children, 2014).

to-child ratio of 1:3. For children older than 3, this ratio is typically 1:6. Children under age 3 attend infant-toddler groups, while those turning 3 move to preschool groups. Variations in group size and age composition are permitted as long as staffing requirements are met. Each center is headed by a principal with an early childhood teacher education degree.

The pedagogical content of childcare centers is rooted in the Nordic tradition emphasizing play-based and child-centered learning as opposed to a more rote instructional approach, common in other educational contexts. All pedagogical content is guided by national guidelines (in the "Framework plan"), which contains the values and purpose of childcare centers, curricular objectives, as well as educational approaches (Ministry of Education, 2006). For example, the "Framework plan" specifies that there should be stimulating verbal and non-verbal interactions in all everyday situations, age-appropriate use of learning materials, and a learning-rich environment that includes work with symbols, books, and reading. Moreover, children spend considerable time playing outdoors year-around.

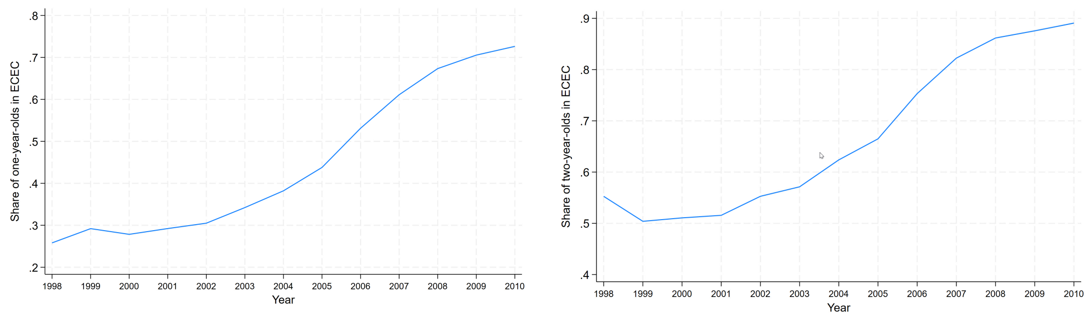
Turning to the school system, children in Norway enter school in August the calendar year they turn 6. Almost all (96%) enter their local public schools, which are free of charge and governed by the municipality. All schools are regulated by the same national curriculum ("Læreplan"), with national maximum requirements for class size and minimum requirements for staffing. Compulsory schooling lasts until 10th grade. We will evaluate the effect of childcare on performance in the school system through 9th grade (usually age 14). In the next section, we describe the reform that expanded childcare in more detail.

3.1 Childcare expansion for 1 and 2 year-old children

In the National Budget for 2001-2002, expansion of childcare coverage was listed as a number one priority for the Government. The stated goal was to reach full coverage (i.e., supply should meet demand) by 2005 (which was not achieved) and on the path towards this establishment of 10,000 new slots in 2003, 4,000 of them for 1- and 2-year-olds. Funding was increased by 10.2 percent from the previous year. The policy was described in detail in a subsequent White paper from the government Ministry of Children and Family Affairs 2003. Municipalities (i.e., local administrations responsible for public childcare) were required to enact the expansion. They received earmarked subsidies for running expenses and additional funding to stimulate expansion (Ministry of Finance, 2021). The National budget signaled that access to childcare would become an individual right from age one by 2006 (whereas in practice it was implemented in 2009), and also signaled an increased focus on quality in 2001-2003. In 2003, an almost unified parliament came to an agreement on the direction of childcare policies which a) maintained the goal of reaching full coverage for 1- and 2- year olds through earmarked funds to the municipalities; b) treated privately and publicly owned centers on the same terms with regard to subsidies and regulations; and c) imposed an income-related maximum parental fee (2750 NOK or USD 400/month, 2004 exchange rate) that was lowered in 2006 (2250 NOK or USD 360/month, 2006 exchange rate). Municipalities had discretion on how the fees varied with household income (Ministry of Children and Family Affairs, 2004).

As described above, the policy reform was enacted in 2003. In Figure 1 we show the expansion of childcare during the 2000s. The expansion is quite flat for the one-year-olds in the years prior to 2003, but increases rapidly in the coming years. For two-year-olds, the increase in the share of slots per child start the year before the childcare agreement in 2002, but we see a sharp increase in the years following 2003. For one- and two-year old children combined, the national coverage rate increased from 46% for children born in 2002 (i.e., first cohort affected by the reform) to 77.5% for children born in 2007, the last cohort for which we have outcome data.

Figure 1: Expansion in childcare for young children



(a) Share of one year-olds in childcare (b) Share of two year-olds in childcare

Source: Statistics Norway

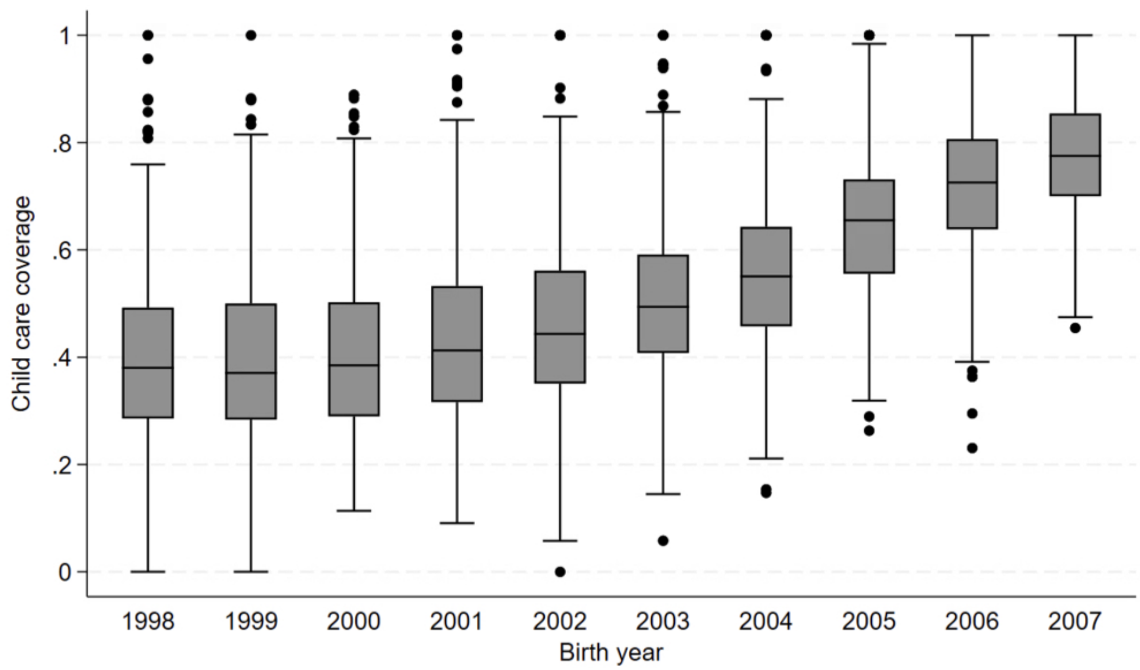
Following the reform, Norway's 428 municipalities were responsible for implementing the scale-up of publicly funded childcare. Progress toward universal access, however, varied considerably across municipalities due to a wide range of idiosyncratic local circumstances. These included unpredictable changes in the number of children born and in demand for public childcare, which in some cases led to concerns about overly rapid expansion and in others to conservative projections of demand; obstacles to new construction, including a lack of available building sites, high construction costs, and shortages of contractors, particularly in the major cities; difficulty recruiting qualified staff, especially in smaller municipalities; and local concerns over the long-term availability of earmarked government funding (see, Eckhoff Andresen & Havnes 2019, Rindfuss et al. 2007, 2010, Asplan Viak 2006, 2009, Zachrisson et al. 2024).[3]

There was substantial variation across municipalities in the childcare coverage across the years of the reform, as shown in Figure 2. Some of the smaller municipalities

[3]These obstacles to expansion were confirmed by those responsible in the Ministry of Education in personal conversations with the authors.

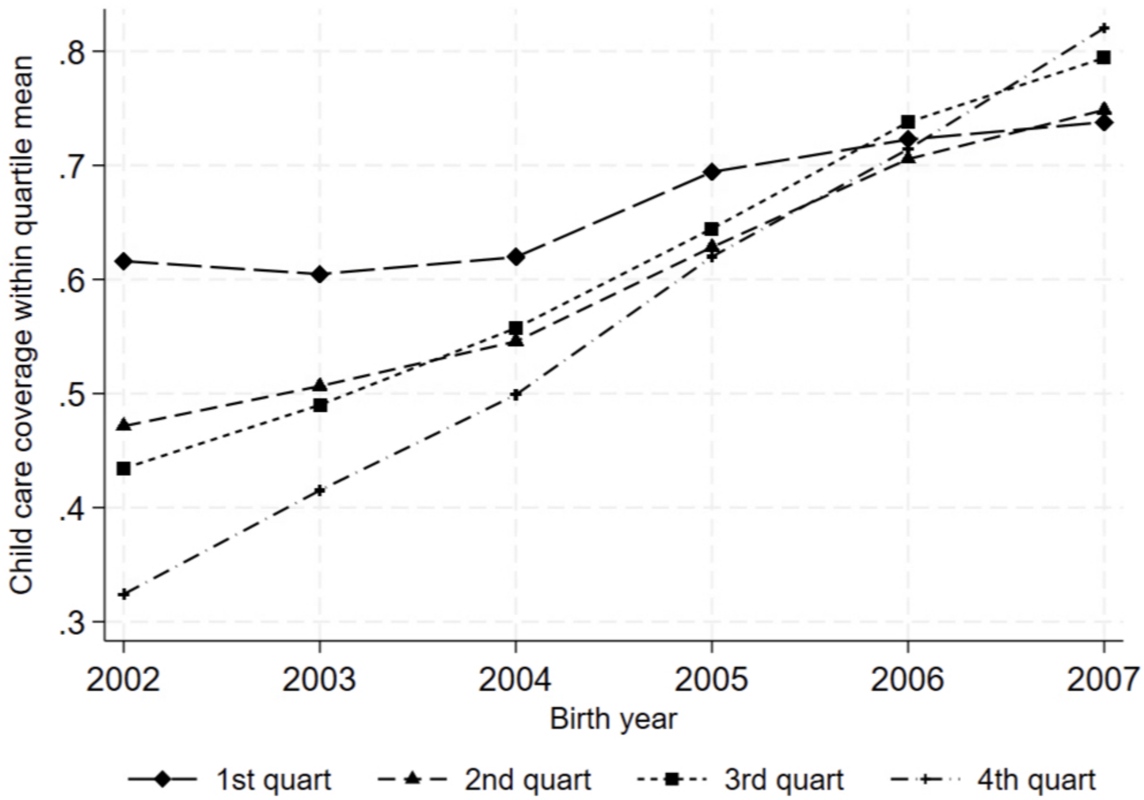
already had almost full coverage for children born in 2002. Throughout the expansion, the variation across municipalities shrank, with the coverage saturating between 70% and 80% for children born in 2007. This is illustrated in Figure 3 , where we graph the coverage rates for municipalities based on quartiles of change rates (created by subtracting the within-municipality coverage for children born in 2007 from that of those born in 2002, and then splitting the rate of change into quartiles.

Figure 2: Variation in Coverage across Municipalities



*Note: The coverage rate reflects the total number of 1- and 2- year olds enrolled in childcare in a municipality divided by the total number of 1- and 2- year olds in that municipality. The horizontal line is the median coverage, the box is the inter-quartile range (IQR), the whisker is +/-1.5*IQR and dots are outliers. Source: Statistics Norway*

Figure 3: Coverage across municipalities by quartiles rates of change



Note: childcare coverage within municipalities (by birth year) by rates of expansion, created by subtracting the within-municipality coverage in 2007 from that in 2002, and then splitting the rate of change into quartiles. The long-dash line is for the quarter of municipalities with the highest initial coverage and lowest rate of change, the dash-dot line is for the quarter of municipalities with the lowest initial coverage and highest rate of change.

4 Data

We start out using data from administrative registries on the entire population of children born in Norway during the years of the expansion of childcare (birth years 2002-2007). These cohorts will be 1-2 years old during the years of the expansion from 2003 to 2008. Children are linked to information about their parents, including level of education, income, and number of children. We link children to municipality of residence at the start of the year when they turn 1. As we rely on the variation in childcare coverage resulting from the expansion, we exclude those without a registered municipality of residence at this point. In Norway, starting childcare prior to turning one is unusual since the vast majority of mothers and fathers will be eligible for parental leave during the child's first year of life.

4.1 National test scores and final exam grades

Our primary outcomes are National test scores in math (numeracy) and reading (Norwegian), taken by all students in Norway in the fall of 5th, 8th, and 9th grade (approximate child ages 10, 13 and 14 years). We focus here on the tests in 5th and 9th grade, as results in 8th grade correlates strongly with 9th grade test results. National tests in Norway are compulsory, while students with special needs and those following introductory language courses may be exempt.

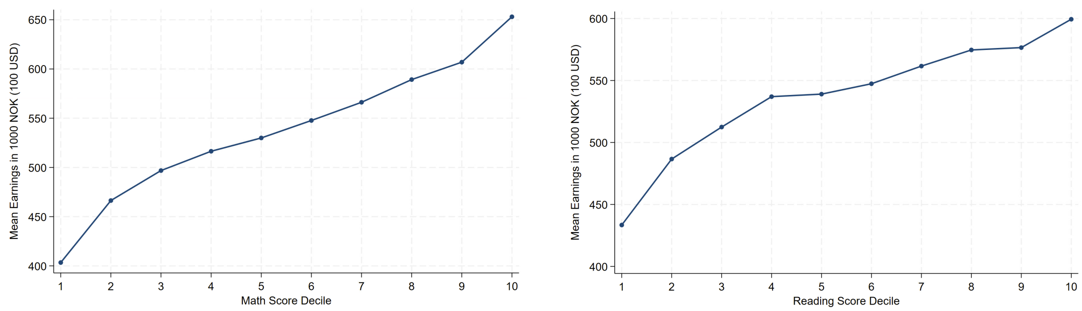
The national tests have been administrated in Norway since 2007. The development of the tests are commissioned by the Directorate of Education and Training, and are constructed by subject experts at universities in Norway along with psychometric experts in the Directorate[4]. The tests are designed to capture the full distribution of skills, and are scored by a national group of teachers blind to school and person identifying information. Identical tests are taken by 8th and 9th graders, to allow estimation of learning across the school year. In the present study, we standardize test scores within year.

In Figure 4, we plot the relationship between standardized test scores in reading and mathematics in eight grade (for the cohort taking the test the first year it was administered, in 2007) and earnings measured at age 29. The figures reveal substantial

[4]Source: *https : //www.udir.no/eksamen − og − prover/prover/nasjonale − prover* downloaded February 12, 2026.

positive correlations: the higher the test score, the higher the earnings. These patterns demonstrate that early academic achievement, as captured by standardized tests, is strongly predictive of later labor market success. This motivates our focus on test scores as economically relevant measures of human capital formation, rather than simply academic indicators.

Figure 4: Correlation between test score decile and earnings at age 29



(a) Math test score decile and earnings (b) Reading test score decile and earnings

Source: Statistics Norway

4.2 Childcare use and municipality level coverage

There is no individual-level registry of childcare use in Norway. Instead, we back out age (by month) of entry into childcare start by using the start and stop date in the cash-for-care register. A family policy at the time, the cash-for-care subsidy was offered to families who did not send their child aged 13-35 months to childcare. A non-taxable cash transfer was offered to families with children not in formal childcare – a replacement if one parent wanted to work less or as a subsidy if buying informal childcare. The register contains the start and end dates for the subsidy receipt for each child separately. The subsidy eligibility stops only if the child starts attending daycare. The acceptance of the subsidy was high after the first years of implementation (Hirch 2010, Arbeids- og velferdsetaten (NAV) 2014).

We identify age of entry into childcare under the assumption that it starts the month after the registered subsidy is terminated. Thus, if the family of receives the subsidy when the child is from 12-19 months, we record the start date to be 20 months. If the family never receives the subsidy, we record the start date as when the child is 13 months, and if the family receives the subsidy until the child is 35 months, when it expires, we record the start date as 36 months. We then validate our age of entry variable by aggregating to a coverage rate by municipality, cohort and age, which we compare to the public record, which gives very similar shares (see Figure 7 in the Appendix). To measure childcare coverage, we lean on data from public records of coverage for children aged 1 and 2 in municipalities by year. Because municipalities fund public childcare, these public records are an accurate representation of the availability of positions, which will be the basis of our instrumental variables approach.

4.3 Individual level control variables

Control variables consist of a dummy capturing the sex of the child and a dummy taking the value 1 if both parents are born abroad to account for immigrant background, as well as controls for birth order and birth month of the child. For parents we construct dummy variables for having finished a high school degree and a college degree respectively, measured the year prior to the birth of the child to avoid endogenous controls. Lastly, we include a control variable capturing age of the mother when the child was born, as well as dummies for the continent of origin of the mother.

5 Empirical strategy

Estimating causal effects of childcare attendance is not straightforward because of selection into childcare by outcome relevant characteristics. In the Norwegian context with high maternal labor force participation, years of childcare attendance is positively correlated with socio-economic status (Drange & Telle 2020). To account for this selection into early care, we instrument for childcare start age using the sharp and staggered increase in coverage rates within municipality by year due to the large increased investment in the wake of the childcare agreement. The increase in funding of new childcare premises in the wake of the agreement, led to a surge in enrollment of young children in childcare across municipalities. Specifically, we use the childcare coverage within municipality by year clusters as an instrument for the age of entry into childcare for a child born the year before (i.e., the coverage rate when the child was one year old (see details of the model specification below). In this strategy, a municipality with a high coverage of childcare slots initially will, typically and mechanically, be subject to a lower increase in coverage. We will pay close attention as to whether these municipalities differ across other dimensions, but note that our strategy relies on changes within a municipality over time.

Formally, our main specification is:

$$\text{First stage: } AGE_{it} = \pi \cdot ChildCareCov_{it} + X_{it}'\beta + \omega_{it} \tag{1}$$

$$\text{Second stage: } y_{it} = \gamma \cdot \widehat{AGE}_{it} + X_{it}'\delta + \epsilon_{it} \tag{2}$$

where AGE_{it} is age in years when child i first attends childcare, $ChildCareCov_{it}$ is coverage of slots in the municipality of residence when the child is 1-2 years old and $\epsilon_{it}, \omega_{it}$ = are residuals. When computing standard errors and p-values we allow for clustering at the municipality level, thus allowing for intracluster and timeseries correlation. X_{mi} are characteristics of the child. In our main specification, we include flexible controls for birth year and year of first placement. Finally, we estimate this regression for subgroups, splitting the sample by the education of the mother, for example, to identify heterogeneity across socio-economic groups.

5.1 Identification

We use the increase in availability of childcare spots as an instrument for attending a childcare center. The instrument is thus the residual of a regression of coverage on time and municipality fixed effects on our observation window. The distribution of this residual is shown in Figure 5 and its support varies from -0.20 to 0.20 showing strong variability.

Standard instrumental variables theory applies when treatment effects are homogeneous across individuals. In this case, IV requires that the instrument strongly predicts treatment take-up and satisfies the exclusion restriction - that is, the instrument affects outcomes only through its effect on treatment participation. When treatment effects are heterogeneous, however, additional identifying assumptions are necessary, and these depend on the specific treatment parameter of interest. We follow the framework developed by Frandsen et al. (2023), who formalize these conditions for the case of judge assignment as an instrument, where different judges exhibit varying propensities to rule in favor of particular outcomes conditional on case characteristics (Kling 2006). While their application focuses on judge instruments, the theoretical framework extends naturally to any setting with a continuous instrument that generates exogenous variation in treatment propensity.

Following their notation, define $\lambda_j = \Pr(J_i = j)$ as the probability that a family lives in municipality j and $p_j = \mathbb{E}[D_i(j)]$ the probability of being assigned to a pre-

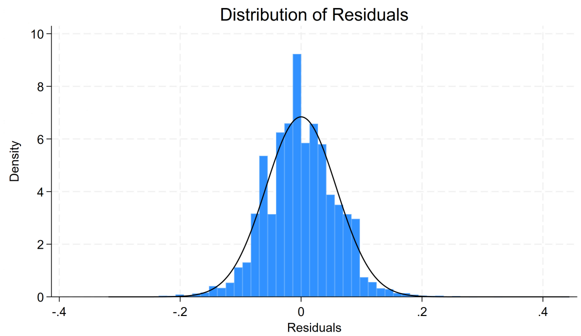


Figure 5: Residual increase in coverage

school, conditional on observed characteristics, municipality and time effects, reflecting the expansion of the program. $D_i(j) \in \{0,1\}$ is the potential treatment when living in municipality j, $\gamma_{ij}(d) = Y_i(d,j) - \bar{Y}_i(d)$ is the effect of treatment d in municipality j while $\gamma_{ij} = \gamma_{ij}(0)\,[1 - D_i(j)] + \gamma_{ij}(1)\,D_i(j)$ is the realized treatment effect.

Monotonicity. With a continuous instrument and heterogeneous impacts we need an equivalent of the monotonicity restriction of Imbens & Angrist (1994). Frandsen et al. (2023) present two versions, the strong or exact monotonicity and a weaker, average monotonicity, which are given below.

$$\text{Exact: } (p_j - p)\big(D_i(j) - \bar{D}_i\big) \geq 0$$

$$\text{Average: } \sum_{j=1}^{J} \lambda_j (p_j - p)\big(D_i(j) - \bar{D}_i\big) \geq 0 \text{ a.s.}$$

The exact version requires that if a child is assigned to a childcare position when a particular expansion takes place they are also going to be assigned to childcare if they were to live in a municipality with a larger expansion. In other words, monotonicity has to hold everywhere on the support of the instrument. Under this condition we can identify an average LATE effect over the whole support, as well as a marginal treatment effect (MTE) over the whole support (Heckman & Vytlacil 2005, Carneiro et al. 2010).

The weaker condition allows identification of the average LATE over the support.

Exclusion. We also need to revisit the exclusion restrictions. Here again, there is a strong version that requires that the impact γ_{ij} does not depend on the municipality j (conditional on municipality and time fixed effects and other covariates) and the weaker version which requires that the conditional covariance of the impacts and the probability of treatment is zero. These conditions are given in the following.

$$\text{Exact: } \gamma_{ij} \text{ does not depend on } j$$

$$\text{Average: } \mathbb{E}\left[\sum_{j=1}^{J} \lambda_j (p_j - p)\gamma_{ij}\right] = 0$$

Risks to Identification. In our context, the exclusion restriction implies that there are no other unaccounted channels for child outcomes that are outcome relevant (affect test scores) and are not orthogonal to the instrument, at least in the weak sense above. This assumption can be restrictive. For example if availability of childcare (our instrument) allows the child to attend and at the same time causes the mother to work, with this having a direct impact on child test scores, beyond the effect of childcare itself, this would invalidate the restriction. The possible reason for an independent effect is increased household resources, which could be used to increase investments in children. To address this risk we present two sets of results. In one, we just use the reduced form where we estimate the impact of childcare availability on child outcomes, without any restrictions on the channels through which this operates. In our next set of results, we implement instrumental variables, which estimates the impact of attendance on compliers (LATE), subject to the validity of the exclusion restriction. In a study of an earlier expansion to the Norwegian childcare system, Havnes & Mogstad (2015) find little effect on female employment and show that formal care crowds out informal arrangements, lending support to the validity of the exclusion restriction.

5.2 Evidence on the validity of the instrument

To provide evidence on instrument validity, we regress childcare coverage on observable family characteristics conditional on municipality and year fixed effects. As shown in Table 1, the coefficients are precisely estimated and very close to zero, with all p-values but three being well above the 5% threshold, and only one (father high school) exceeding the Bonferroni critical value. In all cases, the coefficients are very small. This pattern suggests that the residual variation in coverage is largely orthogonal to observable family characteristics, supporting the exclusion restriction.

Turning now to the predictive power of the instrument for childcare attendance, Figure 6 shows how age of entry has been declining, while coverage for 1-2 year olds has been increasing. The date of the reform, expanding places is marked by the vertical line. The graph shows that immigrants delay entry into childcare, but they are also affected by the reform, albeit more gradually. The data underlying this graph are provided in the Appendix Table 9.

The relevance and strength of the instrument is key to ensuring that the results are asymptotically unbiased (see Anderson & Rubin 1949, Bound et al. 1995, Staiger & Stock 1997, among others). Thus, we begin by evaluating the instrument's relevance overall and in different sub-groups, which we will use for heterogeneity analysis. The first-stage estimates are reported in Table 2. These allow for municipality fixed effects and time effects.

The baseline specification (Column 1) demonstrates that the expansion in childcare availability significantly reduces the child's entry age. A 10 percentage point (pp) increase in coverage for young children in the 1-2 year old range, reduces the average

Table 1: Balancing Test: Instrument Regressed on Control Variables

	Coefficient	t-stat	p-value
Child characteristics			
Female	0.00017	0.850	0.395
Mother's education			
High school	0.00060	1.875	0.061
College	0.00075	1.923	0.054
Father's education			
High school	0.00091	3.370	0.001
College	−0.00022	−0.733	0.463
Other parental characteristics			
Immigrant	0.00153	1.391	0.164
Mother's age	−0.00013	−2.600	0.009
Birth order (ref: First born)			
Second	0.00013	0.542	0.588
Third	0.00048	1.548	0.122
Fourth	0.00136	2.473	0.013
Fifth or higher	−0.00099	−1.269	0.204
Region of origin (ref: Norway)			
Europe	0.00038	0.905	0.366
Asia	0.00075	1.119	0.263
Africa	−0.00037	−0.578	0.563
Americas & Oceania	−0.00007	−0.089	0.929
Constant	0.59565	480.36	0.000

Notes: The dependent variable is coverage (childcare slots for 1–2 year-olds per population of 1–2 year-olds). t-statistics based on standard errors clustered at the municipality. All regressions include municipality and year fixed effects. Birth month dummies (ref: January) are included; individual p-values range from 0.26 to 0.93.

age at entry by about 0.09 of a year (1 month). This is approximately the same across education groups, with a slightly smaller effect for lower educated mothers. The only exception, with a much smaller impact, is for mothers of an immigrant background, where the decline in age of entry is a third as much as for native mothers. In all cases

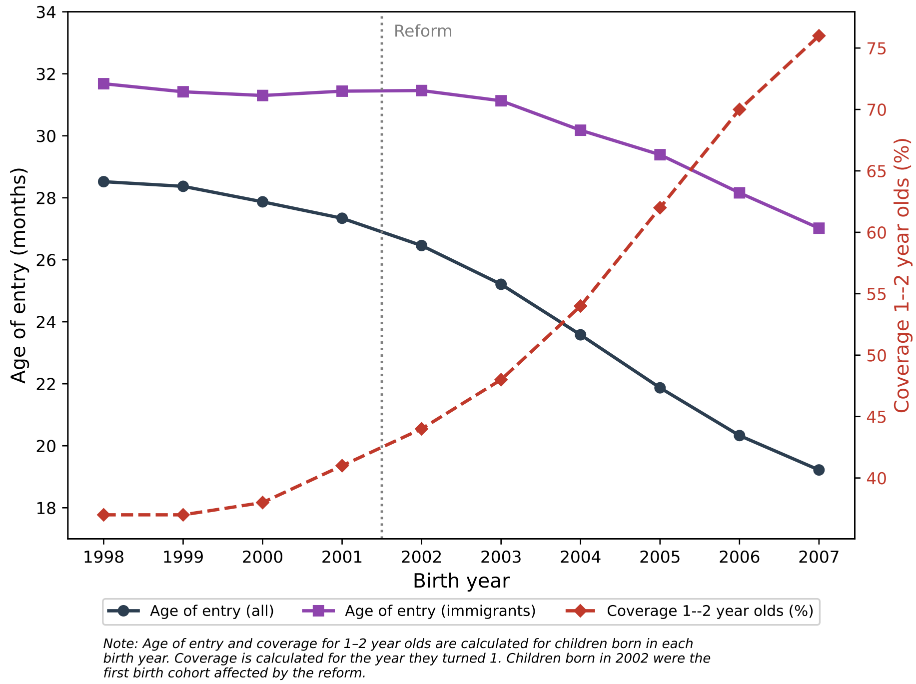


Figure 6: Average age of entry in childcare and Coverage for 1-2 year olds

the instrument is sufficiently strong, even in the case of immigrant families, where the F-statistic is at its lowest, it still is 22.2. In all other cases it is much higher, reaching 149.8 for the sample overall.

Table 2: First stage results

	Impact of the expansion on the Ages at Start of Childcare				
	All	Mum HS	Mum no HS	Immigrant	Native
Coverage	-0.910***	-0.947***	-0.819***	-0.332***	-0.909***
	(0.074)	(0.061)	(0.131)	(0.071)	(0.065)
	[-12.24]	[-15.60]	[-6.26]	[-4.71]	[-13.96]
F-statistic	149.8	243.4	39.2	22.2	194.9
Observations	341,748	254,758	87,035	33,266	308,495

Coverage is the number of places for 1-2 year olds relative to the population of that age group. It can vary from 0 (no places) to 1 (a place of every child). Municipality and year fixed effects are included. Standard errors, clustered at the municipality level, in parentheses. T-statistics in square brackets. HS: High school. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6 Results

6.1 Reduced Form Results

Table 3: Reduced-Form Estimates: Effect of Coverage on Test Scores

	Grade 5		Grade 9	
	Math	Reading	Math	Reading
Coverage	0.069	−0.011	0.087**	−0.014
	(0.046)	(0.040)	(0.044)	(0.044)
Observations	321,899	320,846	314,343	314,948
Municipalities	428	428	428	428

Notes: Coverage is the number of childcare slots for 1–2 year-olds per population of 1–2 year-olds. Standard errors, clustered at the municipality level, are in parentheses. All regressions include municipality and year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Our outcome variables are the test scores from national exams in reading and writing for children at grades 5 and 9. These are scaled to have a unit standard deviation nationally and consequently the results need to be interpreted accordingly. All results include municipality and time effects as well as controls Child sex, immigrant status, dummies for mother and father (separately) having completed secondary and having a tertiary degree, birth month, and birth order.

Overall results We start by presenting the reduced form impact of available places on child mathematics and reading test scores for the entire sample. This "intent-to-treat"-type parameter does not rely on the exclusion restriction, which the IV results presented below assume. Effectively, it allows for channels that may operate in parallel to pre-school attendance, including any direct impact of whether the mother is working on child outcomes, conditional on childcare attendance.

Table 3 implies positive impacts of the availability of childcare on math performance in grades 5 and 9. Mathematics performance increases both in grade 5 and grade 9 and is significant – and robust to multiple testing adjustment – in the latter case. There is no effect on reading performance in either grade, but neither is there a deterioration, a theme we return to later.

Heterogeneity by Education We now explore heterogeneity by education. Table 4 presents the reduced form estimates by education group. The key result here is that increased coverage of early childcare (1-2 years of age) causes a substantial increase in mathematics test scores for children with low education mothers for both grade 5 and grade 9; there are also positive but small and insignificant improvements in reading.

For children with higher educated mothers, there is no effect. The impacts are small and insignificant. However, again, there is no evidence that the availability of childcare harms children.

Finally, note that the Bonferroni critical value for 4 hypotheses and for an overall false positive rate of 5% is 2.5. So, allowing for four hypotheses in grade 5 and four in grade 9, respectively, the improvement in the math test score is significant in both grades 5 and 9.

Impacts on Immigrants In Table 5 we divide the sample by whether the mother is an immigrant or a native. Here again we find a very strong effect of childcare availability on math test scores in grade 9. This would continue to be the case on the basis of a Bonferroni adjustment for four hypotheses (math, reading for natives versus immigrants - critical value for overall 5% is 2.5). This is a result of critical importance in the debate about assimilation of immigrants and bridging the performance gap. We now turn to the instrumental variable results. They help us directly quantify the impact of attendance, which is a key parameter of policy interest.

Table 4: Reduced-Form Effects on Achievement by Parental Education

	Mother no High School		Mother has High School	
	Math	Reading	Math	Reading
Panel A: Grade 5				
Coverage	0.172***	0.054	0.045	−0.019
	(0.065)	(0.064)	(0.048)	(0.042)
Observations	77,283	77,101	244,616	243,745
Panel B: Grade 9				
Coverage	0.230***	0.052	0.053	−0.023
	(0.064)	(0.083)	(0.047)	(0.041)
Observations	74,707	75,021	239,636	239,927
Municipalities	428	428	428	428

Notes: Coverage is the number of childcare slots for 1–2 year-olds per population of 1–2 year-olds. Standard errors, clustered at the municipality level, are in parentheses. All regressions include municipality and year fixed effects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: Reduced-Form Effects on Achievement by Immigration Status

	Immigrant Background		Native Background	
	Math	Reading	Math	Reading
Panel A: Grade 5				
Coverage	0.019	−0.066	0.070	0.016
	(0.106)	(0.105)	(0.046)	(0.041)
Observations	25,954	25,893	295,913	294,922
Panel B: Grade 9				
Coverage	0.293***	−0.014	0.075	0.004
	(0.100)	(0.116)	(0.046)	(0.041)
Observations	25,490	25,543	288,818	289,370
Municipalities	428	428	428	428

Notes: Coverage is the number of childcare slots for 1–2 year-olds per population of 1–2 year-olds. Standard errors, clustered at the municipality level, are in parentheses. All regressions include municipality and year fixed effects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6.1.1 Instrumental Variables Results

Treatment and Inference. The treatment variable is the age of first attendance in childcare. In Norway each family has the option of cash in hand for children kept at home or free childcare attendance. We infer the age of attendance from the time that the family stops receiving the subsidy. Each Table reports the estimated IV coefficient, the bootstrap standard errors clustered at the municipality level, a single hypothesis p-value based on the bootstrap (Efron & Tibshirani 1993, Efron 1987, Cameron et al. 2008), and step-down p-values allowing for multiple hypotheses testing (Romano & Wolf 2005). These correct for multiple testing within grade. So when we look at the overall results, we allow for two hypotheses (reading and math), while when we consider heterogeneity, say by two education groups, we allow for four hypotheses (reading and math for each of the two education groups). We treat grades 5, and 9 separately in this respect.

IV Results. In Table 6 we report IV results for the entire sample split by subject and grades. The coefficients measure the impact of attending preschool a year later. Hence a negative coefficient points to an improvement of test scores for those starting earlier, i.e. at a younger age. For math attending a year earlier improves test scores by about 8% of a standard deviation (SD) at grade 5 and 9.7% SD at grade 9. However, only the grade 9 results is significant at 5%. When we adjust the p-values for multiple testing (within grade mathematics and reading - i.e., two hypotheses) the grade 9 result is still significant at the 10% level. On the other hand the effects on reading scores are effectively zero at both grades and completely insignificant.

These results have two key implications, especially with reference to the results in Baker et al. (2008). Attending preschool earlier, even as an infant (average age at entry post reform varies over time from 26 months to 20 months) improves math performance of 15-16 year olds (grade 9) and while having no effect on reading certainly causes no harm. In what follows we look deeper into this result and how it varies with education.

Table 6: Treatment Effects on Achievement

	Math	Reading
Panel A: Grade 5		
Age at entry	-0.0767	0.0116
	(0.0548)	(0.0486)
Bias-corrected 95% CI	$[-0.185,\ 0.040]$	$[-0.083,\ 0.110]$
p-value	0.152	0.801
Romano–Wolf p-value	0.257	0.801
Observations	321,899	320,846
Panel B: Grade 9		
Age at entry	-0.0974^{**}	0.0151
	(0.0482)	(0.0519)
Bias-corrected 95% CI	$[-0.192,\ -0.003]$	$[-0.088,\ 0.118]$
p-value	0.046	0.754
Romano–Wolf p-value	0.096	0.754
Observations	314,343	314,948

Notes: Age at entry is measured in years and refers to the age at which the child started attending a childcare centre. **The coefficient is the impact on test scores in standard deviation units of attending a year later.** Bootstrap standard errors (1,000 replications) are in parentheses. Control variables include municipality and state fixed effects, and the gender of the child. Bias-corrected confidence intervals and p-values are based on the bootstrap distribution using centered bootstrap t-statistics. Romano–Wolf stepdown p-values control the familywise error rate across both outcomes within each panel, enforcing monotonicity in the stepdown procedure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ based on the single-hypothesis bootstrap p-value.

6.2 IV result heterogeneity

In general, the literature typically finds that the effects of childcare differ between socioeconomic background (Havnes & Mogstad 2011, 2015, Cornelissen et al. 2018, Drange

& Havnes 2019, Gray-Lobe et al. 2022). The impacts may differ because the counterfactual differs: children from high education households will typically offer a richer home environment and will invest more in their children. For them, the advantages of a well-structured early education environment may not be an improvement over the one at home. Indeed, this is the root of the concern that institutionalized care may be detrimental to these children. On the other hand, children from lower SES groups, where the home environment is often not as rich and parenting practices do not follow the best practices (Lareau 2011), may be exposed to a better early education environment, particularly in a context such as the Norwegian, where pre-schools and early care centers follow a strong pedagogical approach with active play and engagement, as well as low child to teacher ratios, as documented earlier.

6.2.1 Results by Education

To investigate whether the effects of early enrollment in childcare vary between different populations, we examine heterogeneous treatment effects by maternal level of education. Table 7 presents IV estimates for these sub-samples, where Panel A reports the results for grade 5, while Panel B those for grade 9.

The results reveal substantial heterogeneity in the effects of early enrollment by maternal education. For children whose mothers lack a high school diploma, we find strong and statistically significant effects showing improvement on math performance across grade levels. Having started childcare a year earlier increases math scores for grade 5 children by 21% of a SD and a p-value of 0.028, while children at grade 9 improve by 28% of SD and a p-value of 0.009. When we adjust these p-values for the fact we are testing four hypotheses in each grade (math and reading mothers with less than high school and more) the significance of these results remains (0.063 for grade 5 and 0.021 for grade 9). Reading test scores also improve, but the results are far from

Table 7: Treatment Effects on Achievement by Parental Education

	Less than High School		High School	
	Math	Reading	Math	Reading
Panel A: Grade 5				
Age at entry	−0.2121**	−0.0675	−0.0402	0.0301
	(0.0845)	(0.0921)	(0.0563)	(0.0482)
Bias-corrected 95% CI	[−0.377, −0.042]	[−0.248, 0.111]	[−0.147, 0.075]	[−0.064, 0.125]
p-value	0.028	0.415	0.440	0.550
Romano–Wolf p-value	0.063	0.786	0.786	0.786
Observations	77,283	77,101	244,616	243,745
Panel B: Grade 9				
Age at entry	−0.2814***	−0.0634	−0.0467	0.0342
	(0.0871)	(0.1222)	(0.0491)	(0.0463)
Bias-corrected 95% CI	[−0.461, −0.115]	[−0.323, 0.156]	[−0.141, 0.046]	[−0.055, 0.128]
p-value	0.009	0.554	0.327	0.436
Romano–Wolf p-value	0.021	0.646	0.615	0.646
Observations	74,707	75,021	239,363	239,927

Notes: Age at entry is measured in years and refers to the age at which the child started attending a childcare center. **The coefficient is the impact on test scores in standard deviation units of attending a year later.** Bootstrap standard errors (1,000 replications) are in parentheses. Control variables include municipality and state fixed effects, the gender of the child, and immigrant status. Bias-corrected confidence intervals and p-values are based on the bootstrap distribution using centered bootstrap t-statistics. Romano–Wolf stepdown p-values control the familywise error rate across all four outcomes within each panel, enforcing monotonicity in the stepdown procedure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ based on the single-hypothesis bootstrap p-value.

significant. When we turn to the children with a higher socio-economic level (mothers with at least high school) we see small and insignificant improvements, in complete contrast to the results for children of lower educated mothers.

Taken together, the results are remarkable. For children from lower SES backgrounds (mother less than high school) attending childcare earlier can be transformational with substantial improvement sin math scores at grade 5 and grade 9. As for children from higher SES backgrounds, although there are no obvious improvements, there are certainly not being harmed by attending preschool, at least in these dimensions we are measuring. This contrasts with the results in Baker et al. (2008) and is crucially important because childcare offers infrastructure that allows women to continue with their career post childbirth. These results show that high quality childcare, as the one offered in Norway, has the potential to improve child cognitive outcomes for children of low SES mothers and does not harm those of higher SES.

Table 8: Treatment Effects on Achievement by Immigration Status

	Non-Immigrant		Immigrant	
	Math	Reading	Math	Reading
Panel A: Grade 5				
Age at entry	−0.0807	−0.0185	−0.1230	0.1218
	(0.0558)	(0.0494)	(0.4993)	(0.4862)
Bias-corrected 95% CI	[−0.188, 0.036]	[−0.112, 0.083]	[−0.975, 1.001]	[−0.746, 1.097]
p-value	0.148	0.721	0.724	0.703
Romano–Wolf p-value	0.335	0.948	0.948	0.948
Observations	295,913	294,922	25,986	25,924
Panel B: Grade 9				
Age at entry	−0.0861*	−0.0061	−0.9174*	−0.0263
	(0.0505)	(0.0484)	(0.8409)	(0.7134)
Bias-corrected 95% CI	[−0.185, 0.014]	[−0.103, 0.092]	[−2.146, 0.020]	[−1.173, 1.019]
p-value	0.087	0.901	0.070	0.935
Romano–Wolf p-value	0.194	0.980	0.363	0.980
Observations	288,818	289,370	25,525	25,578

Notes: Age at entry is measured in years and refers to the age at which the child started attending a childcare center. **The coefficient is the impact on test scores in standard deviation units of attending a year later.** Bootstrap standard errors are in parentheses (1,000 replications for Panel A; 5,000 for Panel B). Control variables include municipality and state fixed effects, the gender of the child, and the education of the mother. Bias-corrected confidence intervals and p-values are based on the bootstrap distribution using centered bootstrap t-statistics. Romano–Wolf stepdown p-values control the familywise error rate across all four outcomes within each panel, enforcing monotonicity in the stepdown procedure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ based on the single-hypothesis bootstrap p-value.

6.2.2 Immigrants

We now consider a separate cut of the data where we compare children whose mother was an immigrant versus those whose mother was a native in Table 8. The results for children of native mothers are similar to those in Table 6. However, those for immigrants show huge impacts: entry to childcare a year earlier causes a more than 90% of a standard deviation improvements in grade 9 mathematics. The result is significant at 7%, although this is no longer the case when we correct for multiple testing. Nevertheless, this result is potentially important because it often the case that children of immigrants under-perform in school. Recall that we estimated a very strong reduced form result for the impact of availability for immigrants – a results that is robust to multiple testing adjustment. Thus, the results suggest that an approach of encouraging early attendance in preschool can improve school performance by immigrant children, helping them bridge the achievement gap with natives.

7 Conclusion

This paper examines the effect of early childcare enrollment on children's academic achievement in Norway, exploiting quasi-random variation in childcare starting age induced by a political compromise leading to an extensive expansion of childcare slots for 1-2 year olds starting in 2003. Using the share of childcare slots available for younger children (ages 1-2) as an instrument for the age at which children start childcare, we find evidence that earlier childcare enrollment improves student performance, particularly in mathematics and for children from disadvantaged backgrounds.

Our instrumental variable estimates indicate that starting childcare one year earlier leads to improvements in test scores of approximately 0.09 standard deviations in mathematics when the child is in grade 9, while we find no improvements in reading scores.

The discrepancy in effects for math and reading may suggest that early childcare in our setting may be particularly beneficial for developing numerical and analytical skills. Moreover, parents tend to be better at emphasizing reading activities and not as well organized and goal driven for activities promoting math with very young children (see Anders et al. 2012, LeFevre et al. 2009, Rege et al. 2024, Karoly et al. 2005, for evidence that home environments tends to support reading rather than math, and on the impact of interventions in math versus reading for slightly older children). Alternatively, the national tests in reading may have less discriminating power at these ages. Moreover, schools in Norway spend considerably more time on reading than math (Directorate of Education and Training 2024), potentially compressing the distribution by the time we measure the effects in Grades 5 and 9.

Importantly, we document substantial heterogeneity in treatment effects across family backgrounds. The benefits of early childcare enrollment are considerably larger for children whose mothers lack a high school education. For this group, we find strong and statistically significant positive effects on mathematics performance across both grade levels tested (grades 5, and 9), with estimates pointing to improvements of 0.28 of a standard deviation at grade 9. In contrast, children of mothers with at least a high school diploma show much smaller and statistically insignificant effects. This pattern suggests that early childcare serves as an equity-enhancing policy intervention, disproportionately benefiting children from less advantaged educational backgrounds who may have fewer learning opportunities at home. Importantly, it also shows that childcare does not cause harm, even to those children from advantaged backgrounds, and hence a much stronger counterfactual. This is in contrast to the well known results in Baker et al. (2008), a difference we attribute to the high quality childcare available in Norway and documented here.

The analysis by immigration background reveals additional nuance, with particularly strong effects emerging in later grades for children from immigrant families, especially in mathematics. Although the IV results are noisy, the reduced form effect of slot availability on mathematics at grade 9 is robust and large. This suggests that early childcare may play an important role in promoting integration and reducing achievement gaps for children from immigrant backgrounds, with benefits that compound over time.

The impact of childcare itself may operate both through the learning opportunities and the interactions with other children within a well-structured environment, or whatever the precise package is implied by “attending childcare”. However, there may also be other channels at play, such as participation in the labor force and family income, although Havnes & Mogstad (2015) did not find these important in Norway. Nevertheless, the reduced form results are subject to fewer identifying assumptions than the IV estimates and reinforce these findings by explicitly showing that slot availability improves outcomes.

Our results align with, while considerably expanding, a line of research on early childcare in Norway (Dearing et al. 2018, Drange & Havnes 2019, Zachrisson et al. 2024, Dearing et al. 2015, Zachrisson et al. 2013) finding positive effects on cognitive and achievement outcomes–primarily for children from low SES backgrounds–and a total absence of negative effects for anyone. When these results are contrasted with the mixed results from studies in other countries (van Huizen & Plantenga 2018), it is worth noting the nature and content of early childcare in Norway. First of all, structural quality is quite tightly regulated, with requirements for teacher training and very low staff:child ratio relative to most other countries. The ”Nordic model” of early childhood pedagogy, emphasizes a holistic, play-based, and child centered approach. When interpreting the effects, it is worth considering which types of skills children acquire during their first years of life which prove beneficial for the fairly complex skills tested in math more than a decade later. Psychologists have recently described what they term FOLD-skills–Foundations of learning development–which comprise a large set of difficult to measure early skills like executive function, critical thinking, curiosity, and creativity (McCoy & Sabol 2025). These skills are hypothesized to be the foundations upon which later skills are built, and to develop in children's day to day play and social interactions. We can therefore hypothesize that the active ingredients of Norwegian childcare centers–child-centered, play-based, holistic–nurture this development. Especially among children from families where such experiences are less common. Regardless, these findings contribute to the ongoing debate about optimal childcare policies, providing evidence that early enrollment can have lasting positive effects on academic achievement that persist through lower secondary school.

From a policy perspective, our findings have important implications for childcare policy beyond Norway. When done right, with high structural quality and age-appropriate curriculum and content, there is a considerable upside for children from low SES backgrounds. While expanding childcare capacity is costly, our results suggest that prioritizing access for younger children, particularly in communities serving disadvantaged families, could yield substantial returns. The particularly strong effects for children of less-educated mothers indicate that early childcare access could be an effective tool for reducing educational inequality and promoting intergenerational mobility.

Our analysis focuses on test score results measured at specific grade levels. Some of the studies discussed above also document effects on non-cognitive skills, and the results for the shorter run are in general mixed – in some contexts, early start may have negative effects on non-cognitive development in early years for some children. Notably, there is no evidence for this in Norway (e.g., Dearing et al. 2015). Long-term effects on other important dimensions of adolescent development, such as social-emotional skills, health outcomes, or–in adulthood–labor market attainment, hence remain important questions for future research.

Despite these limitations, our findings contribute meaningfully to the literature on early childhood interventions and their long-term effects. By documenting both average effects and substantial heterogeneity by family background, we provide evidence that childcare policy can serve as both an instrument for improving educational outcomes and a tool for promoting equality of opportunity and social mobility. The persistence of effects through grade 9 suggests that the benefits of early childcare are not merely short-term, but represent genuine improvements in children's learning trajectories.

References

Almond, D. & Currie, J. (2011), *Human Capital Development before Age Five*, Vol. 4 of *Handbook of Labor Economics*, Elsevier, chapter 15, pp. 1315–1486.
URL: *http://ideas.repec.org/h/eee/labchp/5-15.html*

Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S. & Von Maurice, J. (2012), 'Home and preschool learning environments and their relations to the development of early numeracy skills', *Early childhood research quarterly* **27**(2), 231–244.

Anderson, T. W. & Rubin, H. (1949), 'Estimation of the parameters of a single equation in a complete system of stochastic equations', *Annals of Mathematical Statistics* **20**(1), 46–63.

Arbeids- og velferdsetaten (NAV) (2014), 'Arkiv - kontantstøtte pr. september 2000-2010'. Accessed: 2026-04-14.
URL: *https://www.nav.no/no/nav-og-samfunn/statistikk/familie-statistikk/relatert-informasjon/arkiv-kontantstotte-pr.september-2000-2010*

Asplan Viak (2006), Analyse av barnehagetall per 20.09.2005, Technical report, Asplan Viak, Oslo. Working report, October 2005.

Asplan Viak (2009), Analyse av barnehagestatistikk, Technical report, Asplan Viak, Oslo. Report prepared for the Norwegian Ministry of Education and Research.

Baker, M., Gruber, J. & Milligan, K. (2008), 'Universal child care, maternal labor supply, and family well-being', *The Journal of Political Economy* **116**(4), pp. 709–745.

Baker, M., Gruber, J. & Milligan, K. (2019), 'The long-run impacts of a universal child care program', *American Economic Journal: Economic Policy* **11**(3), 1–26.

Berger, L. M., Panico, L. & Solaz, A. (2021), 'The impact of center-based childcare attendance on early child development: Evidence from the french elfe cohort', *Demography* **58**(2), 419–450.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J. et al. (2017), 'Early childhood development coming of age: science through the life course', *The lancet* **389**(10064), 77–90.

Bound, J., Jaeger, D. A. & Baker, R. M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association* **90**(430), 443–450.

Burchinal, M., Whitaker, A., Jenkins, J., Bailey, D., Watts, T., Duncan, G. & Hart, E. (2024), 'Unsettled science on longer-run effects of early education', *Science* **384**(6695), 506–508.

Cameron, A. C., Gelbach, J. B. & Miller, D. L. (2008), 'Bootstrap-based improvements for inference with clustered errors', *The Review of Economics and Statistics* **90**(3), 414–427.

Carneiro, P., Heckman, J. J. & Vytlacil, E. J. (2010), 'Evaluating marginal policy changes and the average effect of treatment for individuals at the margin', *Econometrica* **78**(1), 377–394.

Cornelissen, T., Dustmann, C., Raute, A. & Schönberg, U. (2018), 'Who benefits from universal child care? estimating marginal returns to early child care attendance', *Journal of Political Economy* **126**(6), 2356–2409.
URL: *https://doi.org/10.1086/699979*

Cunha, F., Heckman, J. J. & Schennach, S. M. (2010), 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica* **78**(3), 883–931.

Currie, J. (2001), 'Early childhood education programs', *Journal of Economic perspectives* **15**(2), 213–238.

Daelmans, B., Darmstadt, G. L., Lombardi, J., Black, M. M., Britto, P. R., Lye, S., Dua, T., Bhutta, Z. A. & Richter, L. M. (2017), 'Early childhood development: the foundation of sustainable development', *The Lancet* **389**(10064), 9–11.

Dearing, E., Zachrisson, H. D., Mykletun, A. & Toppelberg, C. O. (2018), 'Estimating the consequences of norway's national scale-up of early childhood education and care (beginning in infancy) for early language skills', *Aera Open* **4**(1), 2332858418756598.

Dearing, E., Zachrisson, H. D. & Nærde, A. (2015), 'Age of entry into early childhood education and care as a predictor of aggression: Faint and fading associations for young norwegian children', *Psychological Science* **26**(10), 1595–1607. PMID: 26276671.

Directorate of Education and Training (2024), 'Fag- og timefordeling i grunnskolen [subject and time distribution in schools]'. Hentet: 15. mai 2024.
URL: *https://udir.no*

Drange, N. & Havnes, T. (2019), 'Early childcare and cognitive development: Evidence from an assignment lottery', *Journal of Labor Economics* **37**(2), 581–620.
URL: *https://doi.org/10.1086/700193*

Drange, N. & Telle, K. (2020), 'Segregation in a universal child care system: Descriptive findings from norway', *European Sociological Review* **36**, 886–901.

Eckhoff Andresen, M. & Havnes, T. (2019), 'Child care, parental labor supply and tax revenue', *Labour Economics* **61**, 1–19.

Efron, B. (1987), 'Better bootstrap confidence intervals', *Journal of the American Statistical Association* **82**(397), 171–185.

Efron, B. & Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Chapman & Hall, New York.

Felfe, C. & Lalive, R. (2018), 'Does early child care affect children's development?', *Journal of Public Economics* **159**, 33–53.

Fort, M., Ichino, A. & Zanella, G. (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.

Frandsen, B., Lefgren, L. & Leslie, E. (2023), 'Judging judge fixed effects', *American Economic Review* **113**(1), 253–277.

García, J. L. (2024), 'Early education's long-term benefits', *Science* **385**(6715), 1285–1285.

Gray-Lobe, G., Pathak, P. A. & Walters, C. R. (2022), 'The long-term effects of universal preschool in boston*', *The Quarterly Journal of Economics* **138**(1), 363–411.
URL: *https://doi.org/10.1093/qje/qjac036*

Hart, E. R., Bailey, D. H., Luo, S., Sengupta, P. & Watts, T. W. (2024), 'Fadeout and persistence of intervention impacts on social–emotional and cognitive skills in children and adolescents: A meta-analytic review of randomized controlled trials.', *Psychological Bulletin* **150**(10), 1207.

Havnes, T. & Mogstad, M. (2011), 'No child left behind: Subsidized child care and children's long-run outcomes', *American Economic Journal: Economic Policy* **3**(2), 97–129.

Havnes, T. & Mogstad, M. (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. & Vytlacil, E. (2005), 'Structural equations, treatment effects, and econometric policy evaluation', *Econometrica* **73**(3), 669–738.

Hirch, A. (2010), 'Brukere av kontantstøtte 1999-2009 færre barn med kontantstøtte [users of cash for care 1999-2009: Fewer children with cash for care] samfunnsspeilet nr. 4', *SSB* .

Imbens, G. W. & Angrist, J. D. (1994), 'Identification and estimation of local average treatment effects', *Econometrica* **62**(2), 467–75.
URL: *http://ideas.repec.org/a/ecm/emetrp/v62y1994i2p467-75.html*

Japel, C., Tremblay, R. E. & Côté, S. (2005), 'Quality counts: Assessing the quality of daycare services based on the quebec longitudinal study of child development', *Choices* **11**(4), 1–42.

Karoly, L. A., Kilburn, M. R. & Cannon, J. S. (2005), *Early Childhood Interventions: Proven Results, Future Promise*, RAND Corporation, Santa Monica, CA.

Kling, J. R. (2006), 'Incarceration length, employment, and earnings', *American Economic Review* **96**(3), 863–876.

Kottelenberg, M. J. & Lehrer, S. F. (2017), 'Targeted or universal coverage? assessing heterogeneity in the effects of universal child care', *Journal of Labor Economics* **35**(3), 609–653.

Kuehnle, D. & Oberfichtner, M. (2020), 'Does starting universal childcare earlier influence children's skill development?', *Demography* **57**(1), 61–98.

Lareau, A. (2011), *Unequal Childhoods: Class, Race, and Family Life*, 2nd edn, University of California Press, Berkeley, CA.

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D. & Bisanz, J. (2009), 'Home numeracy experiences and children's math performance in the early school years.', *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* **41**(2), 55.

McCoy, D. C. & Sabol, T. J. (2025), 'Overcoming the streetlight effect: Shining light on the foundations of learning and development in early childhood.', *American Psychologist* **80**(2), 135.

Ministry of Children and Family Affairs (2003), 'St.meld. nr. 24 (2002–2003)'.
URL: *https://www.regjeringen.no*

Nelson, C. A., Sullivan, E. & Engelstad, A.-M. (2024), 'Annual research review: Early intervention viewed through the lens of developmental neuroscience', *Journal of Child Psychology and Psychiatry* **65**(4), 435–455.
URL: *https://acamh.onlinelibrary.wiley.com/doi/abs/10.1111/jcpp.13858*

Phillips, D. A. & Shonkoff, J. P. (2000), 'From neurons to neighborhoods: The science of early childhood development'.

Rege, M., Størksen, I., Solli, I. F., Kalil, A., McClelland, M. M., ten Braak, D., Lenes, R., Lunde, S., Breive, S., Carlsen, M., Erfjord, I. & Hundeland, P. S. (2024), 'The effects of a structured curriculum on preschool effectiveness: A field experiment', *The Journal of Human Resources* **59**(2), 576–603.

Rindfuss, R. R., Guilkey, D. K., Morgan, S. P. & Kravdal, Ø. (2010), 'Child-care availability and fertility in Norway', *Population and Development Review* **36**(4), 725–748.

Rindfuss, R. R., Guilkey, D. K., Morgan, S. P., Kravdal, Ø. & Guzzo, K. B. (2007), 'Child care availability and first-birth timing in Norway', *Demography* **44**(2), 345–372.

Romano, J. P. & Wolf, M. (2005), 'Stepwise multiple testing as formalized data snooping', *Econometrica* **73**(4), 1237–1282.

Staiger, D. & Stock, J. H. (1997), 'Instrumental variables regression with weak instruments', *Econometrica* **65**(3), 557–586.

van Huizen, T. & Plantenga, J. (2018), 'Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments', *Economics of Education Review* **66**, 206–222.

Zachrisson, H. D., Dearing, E., Borgen, N. T., Sandsør, A. M. J. & Karoly, L. A. (2024), 'Universal early childhood education and care for toddlers and achievement outcomes in middle childhood', *Journal of Research on Educational Effectiveness* **17**(2), 259–287.

Zachrisson, H. D., Dearing, E., Lekhal, R. & Toppelberg, C. O. (2013), 'Little evidence that time in child care causes externalizing problems during early childhood in norway', *Child development* **84**(4), 1152–1170.

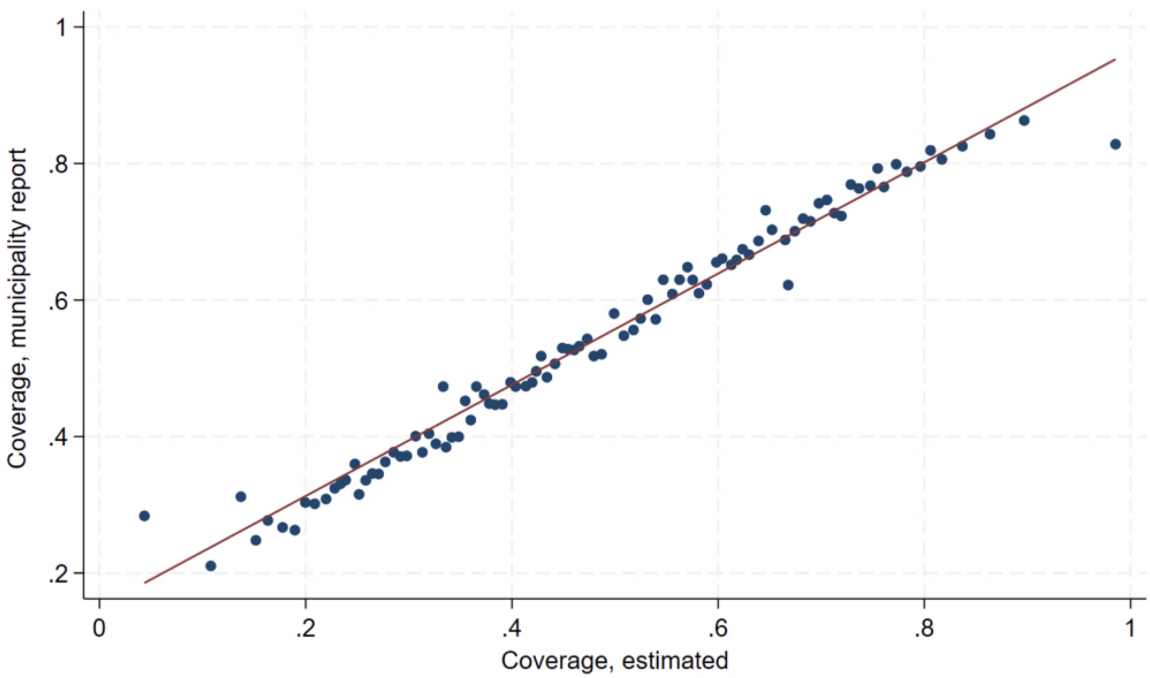
A Supplementary Material

Table 9: Age of entry and coverage rate by birth year

Birth year	Age of entry (all) M(SD)	Age of entry (immigrants) M(SD)	Coverage 1–2 year olds %
Pre reform			
1998	28.52(9.15)	31.68(7.96)	37
1999	28.37(9.18)	31.42(8.17)	37
2000	27.87(9.24)	31.30(8.27)	38
2001	27.34(9.31)	31.44(8.14)	41
Post reform			
2002	26.46(9.44)	31.46(7.97)	44
2003	25.21(9.45)	31.13(8.11)	48
2004	23.58(9.22)	30.18(8.66)	54
2005	21.87(9.10)	29.39(8.86)	62
2006	20.33(8.85)	28.16(9.41)	70
2007	19.22(8.70)	27.02(9.71)	76

Notes: Age of entry is measured in months. Coverage is the number of childcare slots for 1–2 year-olds per population of 1–2 year-olds, calculated for the year the child turned 1. Children born in 2002 were the first birth cohort affected by the reform.

Figure 7: Correlation between municipality reported and estimated childcare coverage



Note: Municipality reported childcare coverage for 1–2 year-olds the year the target child turns one (our instrument) on the Y-axis, coverage estimated from age of entry (our treatment) in that year for on the X-axis (r=.85), across 2002-2007 birth cohorts. The slight misalignment in coverage rates is due to municipality numbers being reported by Dec 31 each year.