Economics of Education Review 98 (2024) 102504

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

The effects of a free universal after-school program on child academic outcomes☆

Nina Drange [a,*], Astrid Marie Jorde Sandsør [b]

[a] Frisch Centre & Statistics Norway, Gaustadalleen 21, 0349 Oslo, Norway
[b] Department of Special Needs Education, University of Oslo, Norway

ARTICLE INFO

JEL classification:
I21
I24
J13

Keywords:
After-school program
After-school care
Difference-in-differences

ABSTRACT

After-school programs provide structured care throughout the academic year. In Norway, after-school programs are an integrated part of school, used to extend the school day to a full working day. Participation is voluntary and is subject to fees paid by parents. In 2016/17, the municipality of Oslo gradually introduced and expanded an offer of free part-time participation in its after-school program, starting with city districts with a high share of children with an immigrant background who were underrepresented in the program at the time. We utilize the staggered rollout of this policy to investigate the effects on enrollment and learning outcomes and conduct an exploratory analysis of student welfare and maternal labor supply. We find that the take-up was substantial, raising enrollment rates from about 70 to 95% in the first wave of affected schools. Despite this, our difference-in-differences estimates show no overall effect of the program on academic performance, neither on average nor across subgroups. There is also little evidence that the program enhanced student well-being or decreased bullying. Heterogeneity analysis does, however, suggest that the program increased earnings for mothers with an immigrant background by about 10 percent.

1. Introduction

After-school programs (ASPs) are structured, adult-supervised programs offered after school throughout the academic year and all day when schools are closed and parents are at work (Seppanen et al., 1993). ASPs facilitate employment and may play a role in enhancing child development and skill formation as several studies show that a lack of adult supervision among school-aged children is associated with lower socio-emotional and academic skills (Aizer, 2004; Bettinger et al., 2014; Blau & Currie, 2006; Martínez & Perticará, 2020; Vandell et al., 2015). The programs could be particularly important for younger children and children from disadvantaged families (Blau & Currie, 2006; Cunha & Heckman, 2007; Kautz et al., 2014). While there is substantial variation across countries, ASPs often include supervision and facilitate group-based activities such as homework time, social interaction, snacks, sports and crafts (Roth et al., 2010; Vandell et al., 2015, 2005).

Studies reviewing the literature on ASPs have shown that participation is typically associated with improvements in both academic and socio-emotional student outcomes (Blau & Currie, 2006; Darling-Hammond et al., 2020; Durlak et al., 2010; Lauer et al., 2006; Levine & Zimmerman, 2010; Vandell et al., 2015). This is also the case for the

☆ We are grateful to Marte Rønning, Nicolai Topstad Borgen, Michelle F. Blair and to participants at the 7th LEER Conference on Education Economics (Leuven), 2022 EffEE Conference on Econometric Evaluation of School Reforms (Munich), the Twelfth International Workshop on Applied Economics of Education (Catanzaro), the Workshop on Educational Resources and Student Performance (Oslo) and the 2023 annual meeting of the Society of the Economics of the Household (Copenhagen). We would also like to thank three anonymous referees for comments and suggestions to earlier versions of the paper. The preparation of this article was supported by funding from the European Research Council Consolidator Grant ERC-CoG-2018 EQOP (grant number 818425) and was partially supported by the Research Council of Norway through its Centres of Excellence scheme, CREATE — Center for Research on Equality in Education (project number 331640).

* Corresponding author.
E-mail address: nina.drange@frisch.uio.no (N. Drange).

[1] Durlak et al. (2010), in a meta-analysis of 68 studies on ASPs conducted between 1980 and 2007, found that program participation predicted increases in self-perceptions, bonding to school, social behaviors, school grades and academic achievement and reductions in problem behaviors. A second meta-analysis of ASPs and summer school programs targeting at-risk students (Lauer et al., 2006), found that participation predicted higher reading and mathematics achievement. Blau and Currie (2006) and Vandell et al. (2015) provide extensive summaries of the research on ASPs and other organized activities in the Handbook of Economics of Education and the Handbook of Child Psychology and Developmental Science, respectively.

https://doi.org/10.1016/j.econedurev.2023.102504
Received 17 February 2023; Received in revised form 17 November 2023; Accepted 7 December 2023
Available online 27 December 2023
0272-7757/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

broader category of organized activities (Vandell et al., 2015), or out-of-school time enrichment activities (Darling-Hammond et al., 2020), including extracurricular activities, sports, community service, drama, summer camps, and other school and community-based programs.[1] However, beneficial effects may hinge on the quality and structure of the program. For instance, Durlak et al. (2010) show that programs that met the SAFE criteria (sequenced, active, focused, and explicit) were associated with much larger effect sizes. Other studies have also found that the benefits of ASPs depend on the quality of the intervention (Kremer et al., 2015; Vandell et al., 2015; Zief et al., 2006). While this research suggests that ASPs can improve student outcomes, many of the studies are correlational and/or focus on specialized programs (e.g., tutoring programs) and associations tend to be modest in size (Vandell et al., 2015).

Causal evidence on the effect of participation in ASPs on child outcomes is scarce. Notable exceptions are Felfe and Zierow (2014), Seidlitz and Zierow (2020) and Schmitz (2022) who study the impact of a universal ASP in Germany on child outcomes. Felfe and Zierow (2014) and Schmitz (2022) find that at-risk students benefit more from ASPs. Felfe and Zierow (2014) employ a value-added approach to show that children of less educated mothers and from low-income families benefit in terms of their socio-behavioral development. Schmitz (2022) uses variation caused by a large reform that subsidized the expansion of ASPs to estimate marginal treatment effects to show that low-SES children have higher ASP premiums. While Seidlitz and Zierow (2020) find positive overall effects on math and language skills using an IV approach, these effects do not differ by SES background.

Additionally, the counterfactual care arrangement has been shown to matter. Using experimental variation in access to a free ASP in Chile, Martínez and Perticará (2020) find no overall effect of the program, but show that students without access to adult supervision at home gain more from the ASP. Such findings can perhaps help explain why some studies find small or no effects of ASP participation, as the comparison group is often comprised of "latchkey children" with no adult supervision, children with supervision from older siblings or other relatives, and children participating in other programs.

In Norway, ASPs are universal and an integrated part of the school, used to extend the school day to a full working day by providing care before and after school and during holidays. In the past decade, increased attention has been paid to the quality and content of these programs and the role they can play in integrating children, particularly children from low-income families or with immigrant backgrounds (Ministry of Children and Families, 2012-2013; Ministry of Education, 2019-2020). As a result, the municipality of Oslo gradually introduced and expanded an offer of free part-time participation in its ASP (denoted in the following as free ASP). The gradual expansion provides us with a unique opportunity to expand on the current causal evidence on ASPs by studying how this cost-reducing policy affects ASP enrollment and, in turn, student outcomes. The program targeted city districts with initial low ASP enrollment and high shares of children with immigrant and disadvantaged backgrounds, allowing us to investigate whether the free program enhanced their skill development, potentially narrowing achievement gaps.

We utilize the staggered rollout in a difference-in-differences analysis to investigate the intention to treat effects of the introduction of the program on (1) enrollment and (2) learning. The project was pre-registered at OSF (https://osf.io/6zrn3) before obtaining the outcome data.[2] Given that many children would have enrolled regardless of the free ASP, we hypothesized that the effects of the program would be concentrated among children from immigrant and low-income families who were more likely not to be enrolled in the absence of the free program. We expected effects to be driven by enrollment at the extensive margin. However, there could also be an effect for children who would have enrolled regardless of the free ASP, either through a change of the peer group, quality and/or content, an increase in enrollment at the intensive margin, or due to reduced costs.

Register data allow us to link children to city districts (and hence treatment status) before they start school, as well as to their families and to test score records from national tests in reading and mathematics. The tests were taken during the fall following four (possible) years of the free ASP. This minimizes attrition and enables a careful analysis of sub-samples by family background, which is important in our setting as many children were already enrolled in the ASP. Focusing on sub-samples allows us to study the outcomes of children in groups that we expect were underrepresented in the ASP before the free program was introduced, such as children from immigrant and low-income families. Going beyond the pre-analysis plan, we use student welfare surveys to investigate the effect of the rollout on measures of well-being and bullying, with individual anonymized data linked to school identifiers. We also use the difference-in-differences specification to analyze maternal labor supply when the child is 6 and 7 years old, in line with previous studies that have investigated effects of ASPs on parental labor supply (e.g., Dehos & Paul, 2023; Felfe et al., 2016).

Our paper contributes to the literature in three important ways. First, causal evidence on the effects of after-school care is scarce, and we make an important contribution to the existing literature by evaluating the effects of a unique ASP reform in Norway. The existing literature is largely focused on one particular reform in Germany (Felfe & Zierow, 2014; Schmitz, 2022; Seidlitz & Zierow, 2020), and we provide new knowledge from a context where participation and the counterfactual differ in meaningful ways. Second, ASPs are regarded as important for social inclusion and potentially also for school achievement, and our findings inform the discussion on universal versus targeted programs for disadvantaged children. Children from disadvantaged backgrounds, including immigrant children, are usually over-represented among those who perform below average in school, and better knowledge about programs and policies that can contribute to "leveling the playing field" is important and of great interest to policymakers. Third, our paper provides evidence of whether a free ASP policy increases enrollment. This knowledge is highly relevant for both ASPs and related programs such as childcare and other organized activities, particularly when the aim is to increase enrollment among disadvantaged children, a group that is often underrepresented in universal programs (see e.g., Drange & Telle, 2020).

Our findings suggest that the take-up was substantial, raising enrollment rates from about 70 to 95% in many affected schools. This shows that the policy achieved the aim of increased enrollment, with close to universal takeup in the years following the new policy. The results, however, suggest little overall effects of the policy on academic performance, neither on average nor across subgroups that were underrepresented in ASPs before the free program and hence more likely to enroll in the ASP following the policy change. There is also little evidence of effects on well-being, bullying, and maternal labor supply.

[2] Our final version differs from the pre-registration plan in the following ways: (1) While we originally pre-registered using all available data, we ended up focusing on cohorts starting school from 2013 and onward, as the trends in outcomes turned out to be very different prior to this year. We do robustness with a longer time series, with results reported in the Appendix. (2) Since it turned out to be difficult to link children to schools before they take the tests in fifth grade, we linked treatment status to the child's city district of residence at the start of the year when they enrolled in first grade. (3) We ended up only studying reading and mathematics, while we pre-registered also including English as an outcome. This was based on the substantial number of outcomes we were already testing, and a realization that reading and mathematics were the most interesting outcomes for our study. (4) We have ended up including three exploratory outcomes, i.e., well-being and bullying and maternal labor supply. These were included based on feedback from conferences and workshops.

Taken together, this suggests that increased enrollment induced by lowering the cost of ASPs did not affect these outcomes. However, with close to universal enrollment, there may be potential to improve these outcomes through quality improvements going forward, although this would imply increased public expenditure.

The rest of the paper is organized as follows: The institutional context and reform details are presented in Section 2, while Section 3 discusses the identification strategy. Section 4 presents the data, Section 5 shows the trends in outcomes and Section 6 presents results. Finally, Section 7 offers a discussion and some concluding remarks.

2. Institutional context and reform details

Organization. In Norway, children start school in August the year they turn 6 years old. The school day for the youngest children typically starts at 8:30 am and ends at 1:30 pm. Before and after the school day, children in first to fourth grade may enroll in the ASP, which most often takes place on school grounds. The ASPs are organized at the municipality level and may be run by both private and public providers, resulting in somewhat varying costs and content across municipalities and to some degree also across schools within a municipality.

With a full-time slot, a child can attend ASPs before school, from 7:30 am, and stay after school ends until 4:30/5 pm. In addition, the child can attend the program during all school holidays, except in July when schools are closed for the summer. With a part-time slot, a child can typically stay in the ASP after school ends and until 3:30/4 pm and can attend the program two days a week during school holidays. Payment depends on family income and many low-income families are eligible for a discount. In 2014/2015, the cost of a full-time slot was about 280 EURO per month for families with a yearly income above 35 000 EURO, 110 EURO for income between 20 000 and 35 000 EURO, and 60 EURO for income below 20 000 EURO. The fee for a part-time slot was 50% of a full-time slot. The majority of children attend the ASP, particularly during their first school years, but the enrollment rate for children from immigrant families has traditionally been lower.

Public primary schools in Norway follow a neighborhood rule with catchment areas determined by the municipality. Most students attend public schools. In 2019, 8 percent of school-aged children in Oslo were not registered in a public school (Rogne et al., 2021). Private schools are mostly religious, with alternative pedagogy or with instruction in other languages such as German, English, or French. Going to a public school other than your neighboring school is not common, but it is possible to apply to and be accepted into other schools with available capacity (Rogne et al., 2021).

Content. ASPs facilitate play, cultural, and leisure activities, adapted to the age, functioning, and interest of the children, as well as provide care and supervision (Ministry of Education, 1998). They are viewed as an important arena for the acquisition of social skills and for language development, particularly among children who speak another language than Norwegian at home. The national framework for ASPs (Ministry of Education, 2020) states their values as recognizing the intrinsic value of childhood, providing security, care, and well-being, being an arena for diversity and inclusion, promoting engagement and joy of creating and the urge to explore, promoting democracy and community, and contributing to sustainable development. The framework is supposed to ensure that all children have access to a program with similar values and content, regardless of where they live.

ASP regulation allows for local variation in organization, content and priority areas and the municipality of Oslo has its own framework for ASPs (Oslo Kommune, 2014, 2018). The framework in place from 2014–2018 included five target areas; (1) nature, technology and the environment, (2) physical activity and play, (3) art, culture and creativity (4) food and health, and (5) homework and in-depth learning (Oslo Kommune, 2014). The framework from 2018 onward includes four target areas; (1) nature, environment, and sustainable development, (2) art culture and creativity, (3) physical activity, and (4) food and health, while pedagogical principles, language development, and basic skill development, including homework help and in-depth learning, are overarching themes. The municipality of Oslo emphasizes learning-supporting activities in their ASP, while other municipalities may focus more on free play and child-initiated activities. This is reflected both in the framework and in leaders' and employees' views on whether learning supporting activities are important, whether ASPs should co-operate with schools about content and whether activities should be child-initiated or should support school activities (Wendelborg et al., 2018).

In an evaluation of the new national framework for ASPs (Caspersen & Utmo, 2022), leaders of ASPs were asked about whether the ASP is an arena where formal language training is carried out. 33% of the leaders responded yes, absolutely, and another 22% responded yes, partly. When asked whether children with minority language backgrounds participate in regular activities, 84% of leaders and 88% of employees responded yes, absolutely, with the vast majority of the remaining responding yes, partly.

In Oslo, parents receive yearly questionnaires about ASPs. The responses from 2018 showed that price was a major reason for not participating in ASPs (Varde Hartmark, 2018). For parents with children in the program, most were satisfied with various aspects of the program. On a scale from 1 to 5, the average score was 4.3 for whether parents experienced that their child enjoyed being in the ASP and was safe, and 4 for whether they were satisfied with the program. When asked whether ASPs contributed to their child's social development or the child's academic development, the average score was 4 and 3.4. While there were no big differences in responses across city districts or grade levels, there were noticeable differences across schools on measures of satisfaction, with the same schools typically scoring high/low across categories, suggesting systematic variation in perceived quality.

The free after-school program. Following an initial trial project at three schools, the city of Oslo gradually introduced a free ASP, beginning with the most disadvantaged city districts and finally including all of the 15 city districts (102 schools). For the municipality, the cost of the program was both covering 50% of the parental payment for a full-time slot, for both previously and newly enrolled children, and an additional municipality-covered cost for each newly enrolled child.[3]

The rollout of the program started in the academic year of 2016/2017 in four city districts (34 schools). Children attending first grade were eligible in 2016/2017, first and second in 2017/2018, first to third in 2018/2019, and first to fourth in 2019/2020. In 2017/2018, the program was expanded to four new city districts (20 schools), with eligibility following the same pattern as the first expansion. In 2018/2019, two more school districts were included in the program (11 schools), but without the same expansion to older grade levels, i.e., only first graders were eligible. Similarly, in 2019/2020, the remaining city districts (37 schools) introduced the program, but only for first graders. In 2022/23 the program was expanded so that all children in second and third grade were eligible for the free ASP. The rollout by school year and grade level is illustrated in Table 1.

The uptake was substantial in schools with initial low rates of enrollment. As illustrated in Fig. 1, enrollment for first graders, measured in October of each year, had a marked increase for all of Oslo in 2016, the first year of the rollout, and continued to increase gradually in the following years as the program expanded. In 2016, rollout increased by about 10 percentage points, from about 80% to 90%, and then increased further to about 93% from 2017–2019 before stabilizing.

Figs. 2(a) and 2(b), showing the take-up for the city districts that introduced the free ASP in 2016/17 and 2017/18, reflect the same

[3] The cost of free full-time enrollment in ASP for grades 1–4 was in 2018 estimated by the Ministry of Education to be about 3000 EURO per student per year for covering parental payments and 930 EURO in additional municipality-covered costs for new slots (Ministry of Education, 2018). For a part-time slot, the estimated cost would be at least half.

Table 1
The rollout of the free after-school program across schools.
Source: Municipality of Oslo.

School year	First grade		Second grade		Third grade		Fourth grade	
	No	Yes	No	Yes	No	Yes	No	Yes
2012/13	102	0	102	0	102	0	102	0
2013/14	100	2	100	2	100	2	100	2
2014/15	99	3	99	3	99	3	99	3
2015/16	99	3	99	3	99	3	99	3
2016/17	68	**34**	99	3	99	3	99	3
2017/18	48	**54**	68	**34**	102	0	102	0
2018/19	37	65	48	**54**	68	**34**	102	0
2019/20	0	102	48	54	48	**54**	68	**34**
2020/21	0	102	48	54	48	54	48	**54**
2021/22	0	102	48	54	48	54	48	**54**
2022/23	0	102	0	102	0	102	48	**54**

Note: The table indicates the number of schools in Oslo of a total of 102 schools without (No) and with (Yes) a free ASP program at each grade level (first to fourth) for each school year from 2012/13 to 2022/23.

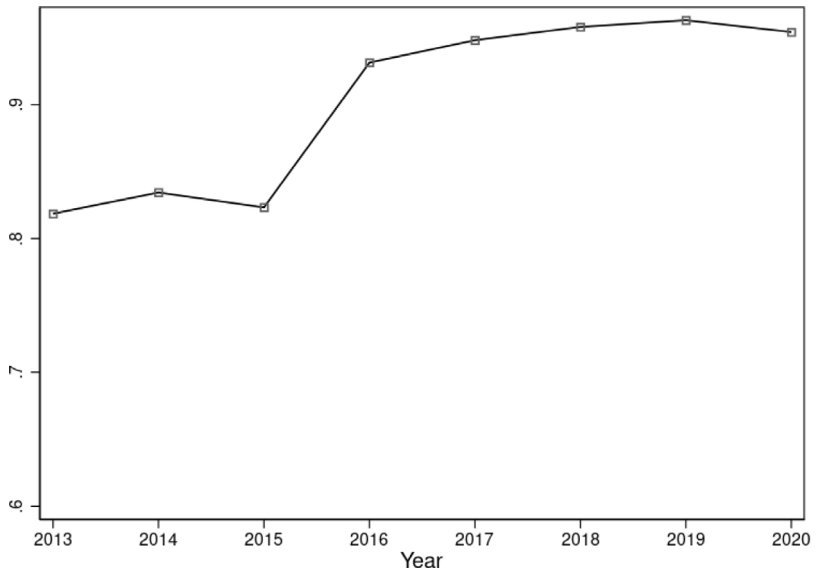


Fig. 1. Share of pupils in after-school programs — entire Oslo, first grade. *Note:* Share of first graders enrolled in the after-school program, measured in October of each year. *Source:* Municipality of Oslo, Department of Education.

pattern found in Fig. 1. In Fig. 2(a), we see that enrollment particularly increased in the four city districts that introduced ASP in 2016/17; Grorud, Alna, Stovner and Søndre Nordstrand. These are all city districts with a relatively high share of immigrants and enrollment in the ASP was initially lower than the municipality average. Enrollment in 2016 increased from 62%–72% to a remarkable 93%–99%, an increase of about 20–25 percentage points. The next school year, the program was expanded to include four more city districts; Gamle Oslo, Grunerløkka, Sagene and Bjerke. For these districts, as seen in Fig. 2(b), initial enrollment in the ASP was higher and there was also a gradual increase in enrollment occurring from the start of the period, 2013, until the year before the implementation, 2016. Still, we see a jump in enrollment of 5–10 percentage points in 2017 when the free ASP was introduced, with enrollment rates stabilizing around 95% thereafter. For the remaining city districts for which we do not yet have outcome measures (with the free ASP for first graders starting in either 2018/19 or 2019/20), initial enrollment was high, already at more than 95% the year before introducing the free ASP, with little potential for further increase as the program was introduced.

Fig. 3(d) illustrates the take-up rate for first to fourth grade in the same period, separate for the city districts that introduced the program in 2016 and 2017 and the city districts that introduced the program later (not yet treated). The figure clearly illustrates the gradual expansion for both treated groups, with increases in enrollment for first grade occurring in the first year of the expansion and subsequently, in the second, third and fourth grade for the second, third, and fourth year of the expansion. The not yet treated group includes city districts that introduced the free ASP for only first graders in 2018/19 or 2019/20 (see Table 1) and for this group, we see very high enrollment rates across the period.[4]

Together, these figures illustrate that the rollout was implemented such that city districts with lower enrollment rates initially, typically city districts with a larger share of low-income families and immigrants (see Table 2), were treated first. Also, the policy was very effective at increasing enrollment in these early districts relative to later districts where enrollment was already high.

It is important to keep in mind that the free ASP consists of two different treatments; (1) changing the counterfactual by bringing kids from other forms of care into formalized ASPs, to a greater extent occurring in early intervention districts, and (2) reducing the cost for families already using ASPs, to a greater extent occurring in later intervention districts. We expect effects to be driven by enrollment at the extensive margin, where students' counterfactual modes of care are important to understand in order to interpret our results. However,

[4] See Table A.3 in the Appendix for estimates of enrollment using the identification strategy described in Eq. (1) of Section 3.

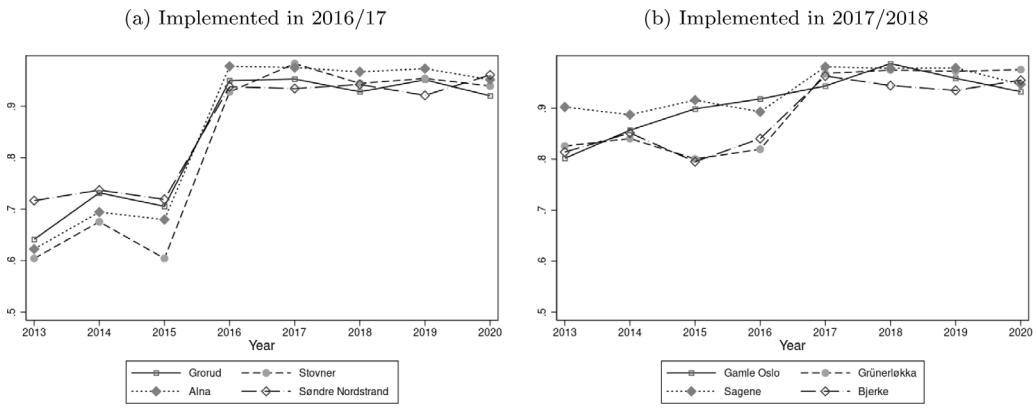


Fig. 2. Share of pupils in after-school programs — treated city districts, first grade.
Note: Share of first graders enrolled in the after-school program, measured in October of each year, covering the period for our main analysis (2013 and onward).
Source: Municipality of Oslo, Department of Education and Statistics Norway.

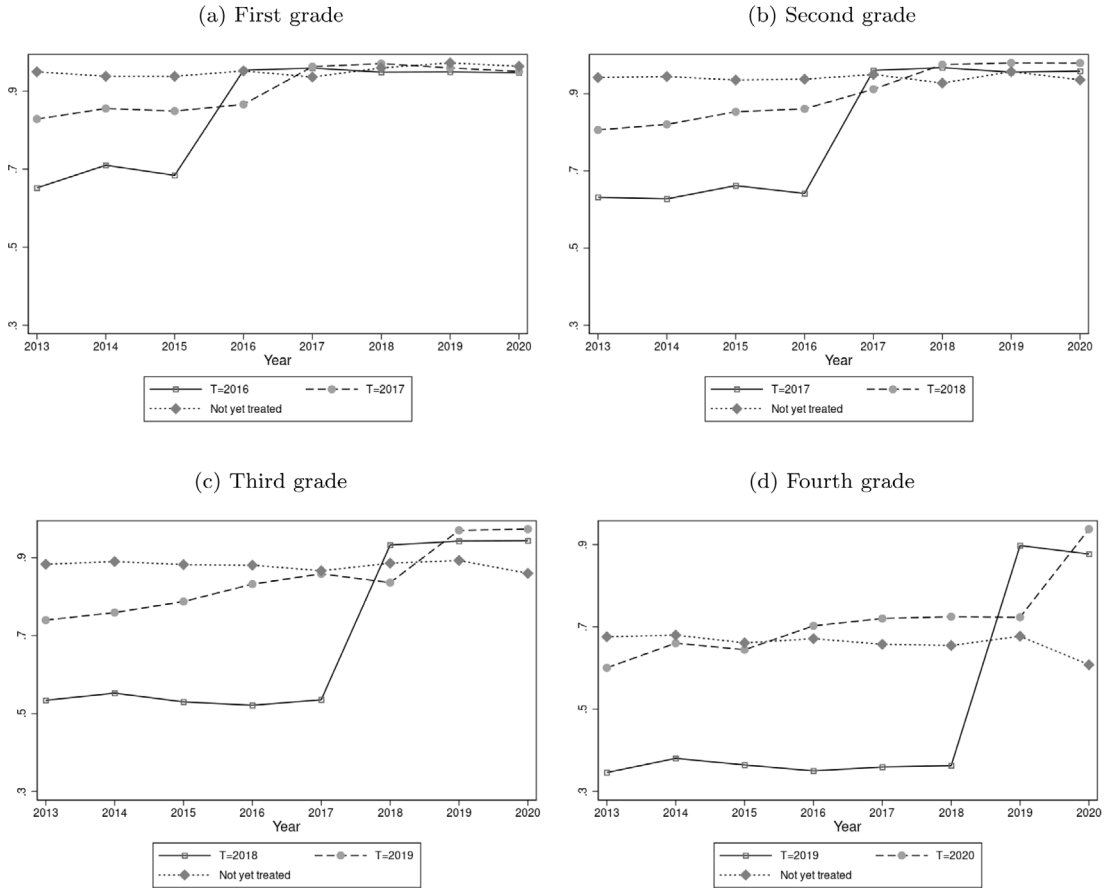


Fig. 3. Share of pupils in after-school programs — by treatment status, first-fourth grade.
Note: Share of first graders enrolled in the after-school program, measured in October of each year, covering the period for our main analysis (2013 and onward).
Source: Municipality of Oslo, Department of Education and Statistics Norway.

there may also be an effect for children who would have enrolled regardless of the availability of the free ASP, either through a change of their peer group, the quality and/or content of the program, an increase in enrollment at the intensive margin or due to reduced costs.

Counterfactual mode of care. An in-depth evaluation was carried out at one of the trial schools, Mortensrud, and included interviews, surveys and register data (PROBA samfunnsanalyse, 2017). This evaluation provides important insights that are relevant for understanding our results. Mortensrud has a particularly large percentage of disadvantaged students. At the time of the trial, 95% of all children at the school had a different language background than Norwegian and only 70% of fathers and 30% of mothers were working. The counterfactual to going to the ASP was mainly reported to be staying at home with mothers in the afternoon (55%), alternatively with other family members or older siblings. Being home with mothers as the main alternative mode of care reflects a high level of parental unemployment at this specific school, particularly for mothers. Many parents also worked part-time and did shift work so mothers/fathers were home in the afternoon. Although alternative care arrangements were available, parents reported wanting children to attend ASPs to participate in activities, learn Norwegian,

Table 2
Summary statistics.
Source: Statistics Norway.

	Treated 2016	Treated 2017	Comparison
Standardized grade language	−0.36	−0.07	0.18
Standardized grade math	−0.34	−0.14	0.19
Proficiency language	0.74	0.82	0.89
Proficiency math	0.72	0.77	0.88
Not taking test Language	0.06	0.06	0.03
Not taking test Math	0.06	0.05	0.03
Girl	0.49	0.50	0.49
Immigrant	0.54	0.36	0.14
Continent of origin			
Asia	0.38	0.18	0.07
Africa	0.13	0.12	0.04
America_Oceania	0.02	0.03	0.02
Europe	0.47	0.67	0.87
Parents background			
Family income	673 840	793 121	1 329 570
Mother has finished high school	0.57	0.72	0.89
Father has finished high school	0.56	0.67	0.86
Mother has a university degree	0.34	0.56	0.75
Father has a university degree	0.30	0.47	0.68
Mother's education is unknown	0.09	0.07	0.03
Father's education is unknown	0.10	0.10	0.05

Note: The first, second, and third columns report the mean of the outcome variables and background variables for children residing in city districts that got access to the free ASP in 2016 (Treated 2016), in 2017 (Treated 2017) and later (Comparison), respectively. The construction of covariates is described in Section 4.3.

and have more friends. The schools receiving the free ASP in our study have on average fewer disadvantaged students, but those who did not previously attend ASPs likely had similar care arrangements, at least in the first wave. Also, note that non-participants in ASPs are more likely to be disadvantaged students (Ministry of Children and Families, 2012-2013; PROBA samfunnsanalyse, 2017) which is quite different from the case of Germany, where low-SES families (low income, low education, single parents, immigrants, and recipients of social transfers) select into ASPs (Schmitz, 2022).

ASP employees. The official school information system (GSI) reports on trained and total staff/student ratios at ASPs. Unfortunately, these data were not collected during the rollout of the free ASP, but there are staff/student ratios available for 2014 and 2021. The data, presented in Table A.4, show that while trained staff/student ratios increased somewhat over these seven years (from about 1 per 50 to 1 per 30 in treated school districts), the staff/student ratio remained constant (about 1 per 10), and this pattern was the same for both treated and untreated districts. This indicates that staff are hired in response to enrollment, regardless of treatment status. The Mortensrud evaluation also provides relevant information about staff composition. Many employees at ASPs work part-time. Out of a total of 12.5 full-time equivalent positions, there were 21 employees, of which 15 were assistants. Assistants typically did not have any formal education beyond upper secondary school although some had upper secondary vocational degrees from relevant fields. This composition of staff is typical for ASPs at other schools in Oslo as well.

3. Identification strategy

The gradual expansion of the free ASP in Oslo implies that a child who started as a first grader in a treated school the year before treatment was not eligible for the free ASP at all, whereas a first grader who enrolled the year after was potentially eligible for four years of the free ASP. This allows us to implement a difference-in-differences model comparing outcomes of children starting school just before and after the program was implemented, across city districts that did and did not implement the free ASP. This strategy will yield unbiased estimates of the ASP if trends in outcomes are similar for children in treatment and comparison city districts, if the composition of families stays similar across districts and time and if treatment effects are homogeneous over time. Our identification strategy mainly relies on a comparison of the results for children attending schools in districts that introduced free ASP in 2016 and 2017 and school districts that did not introduce free ASP until later.[5]

Formally this can be expressed by the following difference-in-differences model with two-way fixed effects:

$$Y_{i,t} = \alpha + \lambda C_t + \beta D_i + \delta DD_{i,t} + \eta X_i^{\tau} + \varepsilon_{i,t} \quad (1)$$

where $Y_{i,t}$ is the standardized result from national tests in fifth grade in reading and mathematics of child i belonging to cohort t.[6] D_i are city district fixed effects indicating the city district that child i resides in the year they start school[7] and C_t are cohort fixed effects. $DD_{i,t}$, our variable of interest, picks up the interaction between residing in a city district that introduces free ASP and belonging to an eligible cohort.[8] X_i^{τ} is a vector of covariates measured at year τ, when the child is age five, and indicates the pre-treatment socioeconomic characteristics of the individual (gender and birth quarter) and parental characteristics (whether mother and father is born abroad, mother's continent of origin, parental education and family income). $\varepsilon_{i,t}$ is the error term with conditional expectation zero. Standard errors are clustered at the city district level, accounting for dependency within the city district.

The model can also be expressed using an event study specification:

$$Y_{i,t} = \alpha + \lambda_t + \sum_{\mu=-7}^{-1} \gamma_\mu D_{i,t} + \sum_{\mu=0}^{1} \delta_\mu D_{i,t} + \eta X_i^{\tau} + \varepsilon_{i,t} \quad (2)$$

where treatment effects are separated into pre-treatment leads (γ) and post-treatment lags (δ) relative to the year before treatment. Finding leads that are not significantly different from the year before treatment lends support to our common trends assumption. Lags that are significantly different from the year before treatment suggest effects of treatment and also show how treatment effects develop over time.[9]

Due to the time between treatment (starting in first grade after the program was introduced) and testing (fifth grade), data is currently available for outcomes in two post-treatment years; for the first and second treated cohorts from the first expansion (test year 2020 and

[5] See Roth et al. (2023) and de Chaisemartin and d'Haultfoeuille (2023) for reviews covering the advances in the recent difference-in-differences literature, including papers by Borusyak et al. (2022), de Chaisemartin and d'Haultfoeuille (2020), Sun and Abraham (2021), Callaway and Sant'Anna (2021) and Goodman-Bacon (2021).

[6] In the pre-registration plan, we also suggested studying results on national tests in English, but we concluded that the Norwegian reading test should be sufficient to pick up changes in language proficiency.

[7] We assign children to city districts in the start of the year when they turn 6 years old, i.e., about eight months before they start school.

[8] For the case with a single treatment period, the model can be expressed as $Y_{i,t} = \alpha + \lambda_t + \beta D_i + \lambda(D_i \times Post_t) + \eta X_i^{\tau} + \varepsilon_{i,t}$ where D_i is a dummy variable equal to 1 if the child lives in a city district that introduced the free ASP during the period we study and λ_t are cohort fixed effects that absorb the post-treatment indicator. $D_i \times Post_t$, the variable of interest, is a dummy variable equal to 1 if the child lives in the treatment area and starts school in or after the year the free ASP was introduced.

[9] This interpretation assumes homogeneous treatment effect profiles, as estimates for one relative period are potentially contaminated by the effects of other relative time periods in the sample, including the excluded time period (Sun & Abraham, 2021). For our event study specification, we present results using the Callaway & Sant'Anna estimator csdid (Callaway & Sant'Anna, 2021; Rios-Avila et al., 2022), which takes this into account.

2021) and for the first cohort from the second expansion (test year 2021). For pre-years, we initially pre-registered that we would use all cohorts available, which in practice would imply children starting school in 2010 and onward.[10] However, after a careful inspection of the trends in the pre-treatment period, our main analysis focuses on children starting school from 2013 and onward as there seems to be a trend break in test scores for all groups between 2012 and 2013. We report results from analyses where we also include cohorts of school starters 2010–2012 as a robustness check. Lastly, we need to pay particular attention to the three trial schools where the program was implemented before the rollout. In the analysis, we exclude pupils who take their national test at these schools.

With multiple periods and variation in treatment timing, heterogeneous treatment effects may cause biased estimates due to early adopters entering the control group. For our time window, the worry is that the first expansion group (early adopter) becomes a control group for the second expansion group (later adopter). We investigate the potential for these biases by presenting the results from an event study specification, as well as presenting results using the Callaway & Sant'Anna estimator csdid (Callaway & Sant'Anna, 2021; Rios-Avila et al., 2022). When we estimate the results using csdid, we specify the not yet treated option. This allows us to use the not yet treated city districts as a control for those receiving the free ASP in the first wave of the rollout.

In our pre-registration, we only included academic outcomes. However, we have added an exploratory analysis where we investigate whether the expansion of ASPs affected maternal labor supply. We estimate Eq. (1) for two separate outcomes: mothers' linear earnings in the year the child turns 6 and the year the child turns 7, respectively. In an additional exploratory analysis, we use data from a student survey that is run every year by the Directorate of Education where fifth grade students answer questions about their experience of being in school and the school's social environment. We study the two topics covered in the survey that we think hold relevance for our purpose, bullying and well-being, and estimate effects using Eq. (1).

4. Data

4.1. Sample construction

Our study population is primary school students in Oslo and their families, collected from the demographic registries of Statistics Norway. We include all children registered in a city district in Oslo on January 1 the year when they are eligible to start school. Children are linked to parents and siblings with a unique identifier. For children, we include data on educational outcomes and information about school district, gender, and immigrant background. For parents, we include information on family income and educational attainment, as well as the mother's continent of origin.

The rollout of the program is linked by school identifiers and year. The data on the rollout of the free program was collected from the municipality of Oslo and includes information on the treatment status for each school and grade level each year. We include test scores for children starting school from 2013 (2010 in robustness analyses) and onward.

Three schools in two different city districts were part of a trial project where they had access to the free ASP before the rollout. In these schools, there was also an emphasis on the content of the program. National test results from these schools were excluded from the analysis, however, results change very little if these schools are included.

To study the effects of the free ASP on student welfare, we use data from a student welfare survey called Elevundersøkelsen that takes place every year in the fall of fifth grade.[11] We have individual data linked to school identifiers, but these are anonymized and cannot be linked to the rest of the data at the individual level, preventing us from studying sub-samples for this outcome. We have access to tests from 2017–2021 and thus our data cover the same cohorts as in the main analysis, i.e., children starting school from 2013–2017.

4.2. Outcome variables

The main outcome variables are collected from compulsory national tests in the subjects of reading and mathematics. These tests are taken at the start of fifth grade, i.e., the cohort starting school in 2013 and tested in the fall of 2017. The outcome consists of a continuous variable that measures overall performance as a scaled score based on a 2-parameter item response theory (IRT) model, as well as a categorical variable taking the value 1, 2, or 3 depending on performance. The tests are developed and validated by experts in test development and psychometrics and are designed to capture the full range of skills in these subjects. The results are mainly used to collect information about students' skills and to track school development over time. Results are conveyed to teachers and parents but have no direct consequence for students apart from the aim of adapted education. About 96% of all students in Norway take the test; students with special needs and those following introductory language courses may be exempted. Note that the register data allow us to follow children who move from Oslo during or after treatment. Since the national tests are taken across the country, a child who lives in a city district in Oslo before starting school (and is registered in our sample), may move to another part of the country, but still be registered with academic outcomes in our data.[12]

We show separate estimates for reading and mathematics using both the scaled score as well as dummy variables for each subject that measure if the individual scores 2 or 3 (and not 1). The latter variable is constructed to better capture changes in reading or math proficiency among low-performing children. In addition to the test scores outcomes, we construct dummy variables that capture if the child was exempted from the test.

To explore whether the free ASP program induced some mothers to work more, we construct an earnings measure capturing mothers' linear earnings[13] the year their child or children in the sample turn 6 and 7 respectively. We emphasize that because this outcome was not pre-registered in our data, this analysis should be regarded as exploratory.

In another exploratory analysis, we use the student welfare survey to look at potential effects on student well-being and perceived bullying. The well-being measure is based on one single question: "Do you

[10] Due to a child care program providing free part-time child care slots in the same city districts that received the first wave of ASP treatment, including cohorts starting school prior to 2010 would imply that post and pre-cohorts could have differing trends due to other reasons than the ASP policy.

[11] The Student Survey can be carried out in all grades from fifth and onward. The survey must be carried out in 7th, 10th, and 11th grade, whereas it is decided on the school level whether the survey is carried out in other grades as well. Students can choose whether to answer. Our sample size is reduced from about 35 000 to about 28 000 when we go from looking at school academic outcomes to student welfare. The number of schools is very similar and the reduction in sample size is mainly caused by non-response among students. As we are not able to link these data to register data, we are unfortunately not able to investigate the characteristics of non-responders.

[12] Overall, the share of children that we find registered in a city district in Oslo on January 1 the year they start school and who take the national test in another municipality in fifth grade is 11.6%.

[13] We adjust income by the basic amount (Grunnbel≠pet), which is adjusted for inflation and wage growth by the Norwegian parliament every year and is included in most formulas for public welfare transfers.

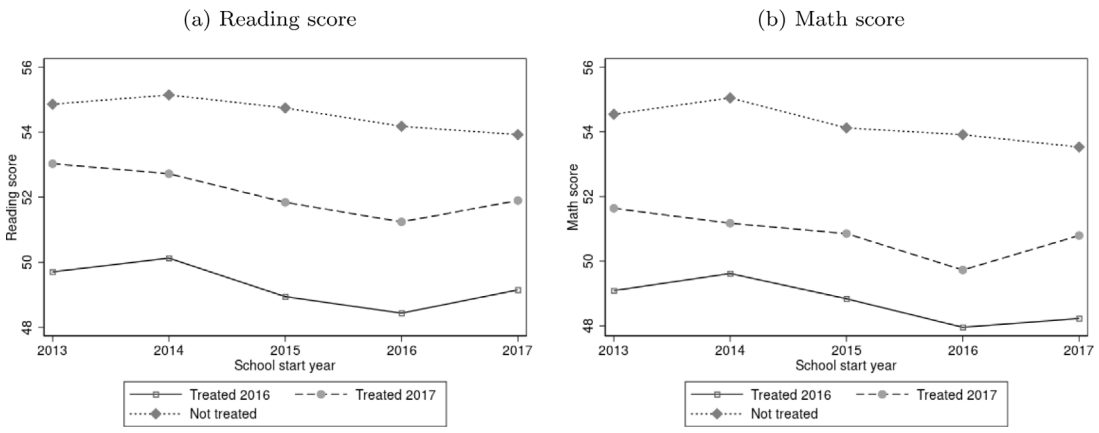


Fig. 4. Trend in national test scores, all students.
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

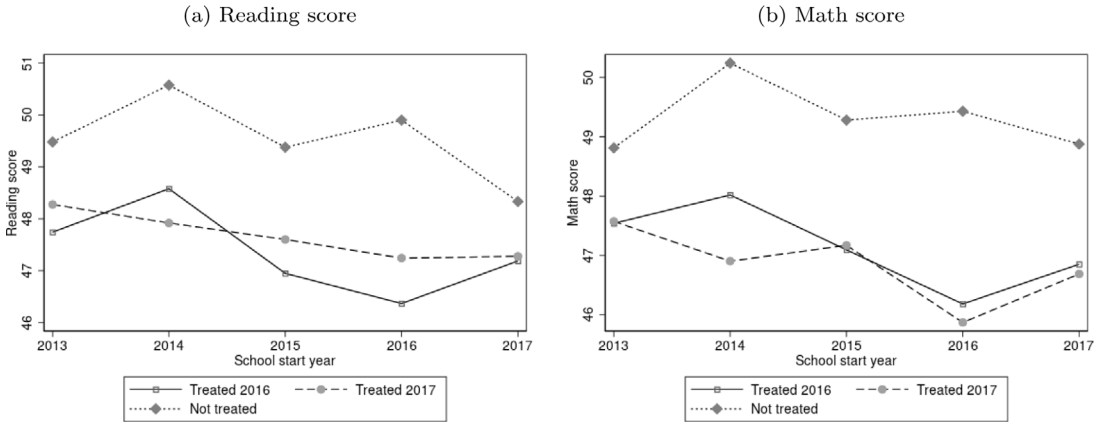


Fig. 5. Trend in national test scores, immigrants.
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

enjoy school?"[14] and the answer is graded from 5 (Yes, absolutely), to 1 (Not at all). The bullying outcome is based on the row-mean of two questions, where the first is "Have you experienced bullying in school during the previous months?" and the second is "Have you experienced digital bullying (on cell phone, iPad or PC) during the previous months?".[15] The answer is graded from 5 (Not at all) to 1 (Several times a week).

4.3. Control variables and sub-sample stratification

The background characteristics of the children and their families, including family income and education, are measured the year before the child starts school to ensure that they are not endogenous to treatment. For the child, we construct a dummy taking the value 1 if the child is female and 0 if male, as well as dummies for quarter of birth. Having an immigrant background is defined as having two parents born abroad. We also construct dummies for the mother's continent of origin. The control for family income is the average income of the mother and father. For educational attainment, we construct dummies on whether the mother/father has finished high school or college, respectively. Following the pre-registration plan, we study sub-samples by mother's education (finished high school or not), immigrant background, and whether family income is below 60% of median family income.[16]

Table 2 shows descriptive statistics separately for both treatment groups and for the comparison group. The 2016 treatment group has a higher share of immigrants (0.54) than the 2017 treatment group (0.36) and both are higher than the comparison group (0.14). The share with an immigrant background from Asia and Africa (mother's continent of birth) is 0.38 and 0.13 in the 2016 treatment group, 0.18 and 0.12 in the 2017 treatment group and only 0.07 and 0.04 in the comparison group. Parental education is gradually increasing from one group to the next. As pointed out earlier, this is expected since the rollout started with the city districts with the lowest socioeconomic backgrounds and continued to the city districts next in line. The identifying strategy when using difference-in-differences, however, hinges on common trends rather than common levels as well as the composition of families staying similar across treated and non-treated districts.

5. Trends in outcomes

We begin with a visual inspection of trends in our outcome variables separately for treatment and control groups. Figs. 4(a) and 4(b) show

[14] Original question in Norwegian: "Trives du på skolen?"

[15] Original questions in Norwegian are respectively: "Er du blitt mobbet av andre elever på skolen de siste månedene?" and "Er du blitt mobbet digitalt (mobil, iPad, PC) de siste månedene?"

[16] By mistake, we pre-registered that low income should capture children with a family income below 40% of the median. The intention was to use the EU60 measure, which is what we have done in the analysis.

Table 3
Main results.
Source: Statistics Norway.

	Reading			Mathematics		
	Score	Proficiency	Exempted	Score	Proficiency	Exempted
Panel 1: Two way fixed effects						
$D^{2016/2017}$	0.017	−0.009	−0.002	0.017	0.004	−0.001
	(0.037)	(0.008)	(0.005)	(0.028)	(0.008)	(0.004)
CI with wild bootstrap	[−.075, .104]	[−.028, .010]	[−.014, .010]	[−.051, .086]	[−.015, .022]	[−.010, .008]
N	33079	33079	34659	33086	33086	35026
Panel 2: Two way fixed effects with covariates						
$D^{2016/2017}$	0.006	−0.013	−0.001	0.010	0.002	−0.000
	(0.035)	(0.008)	(0.005)	(0.026)	(0.009)	(0.004)
CI with wild bootstrap	[−.078, .084]	[−.033, .005]	[−.013, .011]	[−.053, .077]	[−.018, .020]	[−.009, .009]
N	33079	33079	34659	33086	33086	35026
Panel 3: Callaway & Sant'Anna (csdid)						
ATT	0.067+	0	0.003	0.021	0.007	−0.001
	(0.041)	(0.010)	(0.005)	(0.036)	(0.013)	(0.007)
CI with wild bootstrap	[−.005, .139]	[−.020, .020]	[−.006, .013]	[−.054, .096]	[−.015, .012]	[−.020, .034]
N	33079	33079	34659	33086	33086	35026
Mean	50.3	.78	.06	49.8	.75	.06

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Panel 1 reports the results based on Eq. 1, whereas Panel 2 displays outcomes from Eq. 1 with covariates described in Section 4.3. Standard errors in both are clustered at the city district level, whereas the CI interval reported is calculated from standard errors using the wild bootstrap procedure. Panel 3 reports results from the csdid estimator described in Callaway and Sant'Anna (2021), with standard errors estimated with the wild bootstrap procedure (wboot). Includes cohorts starting school in Oslo from 2013–2017. Mean refers to children in city districts that received treatment in 2016 or 2017, measured in 2015. In the regressions, the score is standardized, and hence the estimates are reported as percent of a standard deviation.

the average national scaled test score for reading and mathematics, respectively, for cohorts starting first grade from 2013 to 2017, separately for the 2016 and 2017 treatment groups and the comparison group. Although scores in reading and mathematics are both higher in comparison than treated districts, as expected, the pattern in trend is the same for both treatment and control districts for the three years before treatment,2013–2015, while the trends are less stable before 2013.

For the 2016 treatment group, there are two cohorts in the post-treatment period. For both reading and mathematics, test scores in the first cohort (started school in 2016) relative to the comparison group do not seem to change, but there is a slight relative increase for the second cohort (started school in 2017). For the 2017 treatment group, we see an increase for the first treated cohort (who started school in 2017) relative to the comparison group.

A second notable observation from the figures is the comparison of the two treatment groups in 2016, as this is a comparison of early and later treated units. As the reform was gradually rolled out to less and less socioeconomically disadvantaged school districts, these are the most comparable in terms of background characteristics. In both figures, we see no sign that scores developed differently for these two groups in 2016.

However, average effects could conceal important heterogeneity. Specifically, we hypothesized that children with immigrant backgrounds would benefit the most from the policy, both due to their low relative enrollment before the introduction of the free ASP and because enrollment would increase their daily exposure to a setting where the Norwegian language is dominant. If this was indeed the case, we would expect stronger effects for reading than for mathematics. In Figs. 5(a) and 5(b) we therefore look at the same trends for the immigrant population only. The average scores are somewhat more jumpy in the pre-treatment period than for the population as a whole, reflecting the now smaller sample.

There does not seem to be any visual evidence of a positive effect on school performance of the free ASP for the first cohort of treated first graders (2016 for the 2016 treatment group), as the treatment group does not show any growth for treated children relative to the comparison group. For 2017 there is some sign of an improvement for both treatment groups relative to the comparison group.

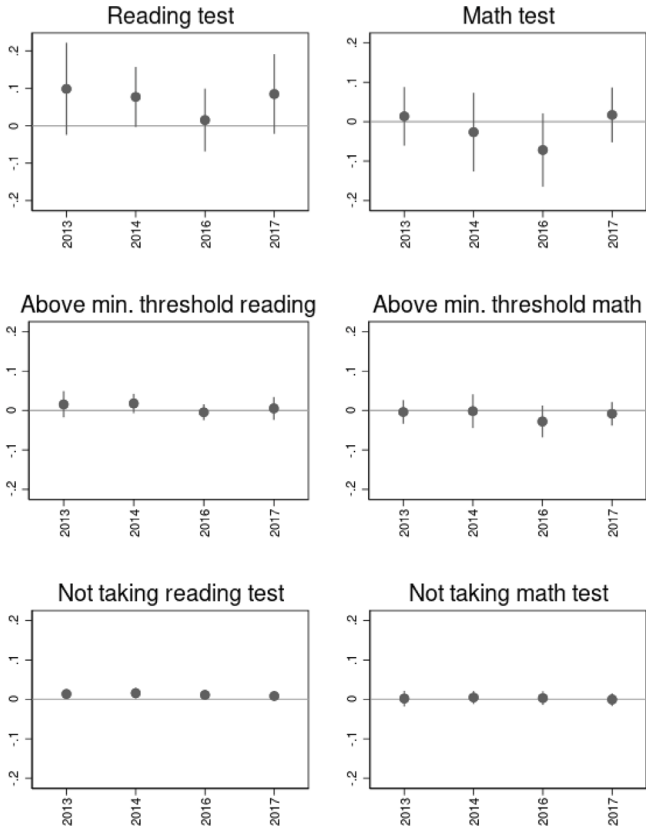


Fig. 6. Event study estimates of student outcomes, all children
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

To investigate whether there is heterogeneity in the distribution of effects, we also look at a specification using an indicator for where reading and mathematics scores are above a lower threshold, i.e., whether the categorical proficiency level is 2 or 3 rather than 1 (see Figs. A.1(a) and A.1(b) for the whole population and Figs. A.2(a) and A.2(b) for children with immigrant backgrounds in the Appendix). We see that the trends follow each other more closely over the pre-treatment period. There is still no visual sign of an effect of the free ASP on the first cohort of treated children (2016) while there is a slight relative increase in the second year (2017) where both treatment groups have treated children.

One reason to exempt children from the tests, would be if their language proficiency is so poor that it becomes difficult to understand the test questions. This implies that being exempted from the test is an outcome in itself, as treatment could potentially affect this margin. If more children take the test because their language proficiency has improved, we might also see a decrease in the mean results, as these children likely belong to the lower end of the distribution. We investigate trends for whether students are exempt from the test (see Figs. A.3(a) and A.3(b) for the whole population and Figs. A.4(a) and A.4(b) for children with an immigrant background in the Appendix). Here the trends are less stable in the pre-treatment period. In the post-treatment period, there was a slight decrease in the share not taking the test in 2016 for the treatment group relative to the comparison group, followed by a small increase in 2017.[17]

6. Results

6.1. Main results

In Panel 1 and Panel 2 of Table 3, we present results based on an estimation of Eq. (1), with and without covariates. Standard errors clustered on the city district level are in parentheses, whereas confidence intervals for the estimates are reported with wild bootstrapped standard errors. Outcome variables are scaled scores, normalized to have a mean of 0 and standard deviation of 1 for the entire period, an indicator for proficiency and whether students are exempted from the test, for both reading and mathematics. We restrict the estimation window to the cohorts starting first grade in the years 2013–2017, where we observed stable pre-trends across outcomes.[18] In light of the new developments in the difference-in-differences literature, Panel 3 presents results from the Callaway & Sant'Anna estimator csdid.[19] The results reflect what we observed when inspecting the trend figures — for most outcomes, there are no measurable effects of introducing the free ASP on national tests, regardless of specification. The signs of the estimates are in the expected direction, except for reading proficiency in the two-way fixed effects specification. Adding covariates in Panel 2 barely moves the estimates. The 95% confidence interval for the reading result is [−.078, .084], which implies that we can reject negative and positive effects larger than 8% of a standard deviation. We note that this confidence interval includes quite large effect sizes, in both a negative and positive direction. In Panel 3, where we show results from the specification that implements the csdid estimator, all estimates are either zero or have a positive sign. We see a positive estimate of 6.7% of a standard deviation for reading, significant at the 10% level, whereas the other estimates are mostly small and not significant. The 95% confidence interval for the reading result in the csdid specification is [−.01, .15]. From the second column, we note that the estimate for reading proficiency is precisely estimated in this specification. This indicates that the positive estimate for reading is likely not driven by students in the lower part of the distribution.

As mentioned previously, average effects could conceal important heterogeneity. We are particularly interested in whether children with immigrant backgrounds benefit from the policy, in addition to children from families with low income or where the mother is a high school dropout. We should keep in mind that a motivation behind the policy was to provide children from immigrant families with the opportunity to enroll in the ASP. Hence, even though we do not observe individual enrollment, we expect that this is the sub-group that likely is more directly affected by the policy. At the same time, the change in peer composition and the increased number of children could potentially change the dynamics of the ASP environment for children that were already enrolled, although the expected direction and size of this effect is not necessarily clear (Epple & Romano, 2011; Leuven & Oosterbeek, 2018; Sacerdote, 2011; Schanzenbach, 2020).

Results using the two-way fixed effects estimator without covariates are presented in Table 4 and show few signs of heterogeneous effects.[20] If anything, the signs are in the opposite direction of what we would expect, with positive signs for children from a more advantaged background. In the Appendix, Table A.1, we show similar results with covariates included. It is reassuring that the results are robust to this inclusion. Note that while the lack of effects suggests that the program did not promote language learning in the population of children with an immigrant background, it may be seen as reassuring that results for the population likely already enrolled are unchanged. This indicates that the inflow of children from less advantaged backgrounds to the ASPs did not have a negative spill-over effect on those who enrolled regardless of the policy.

Next, we present results for the event study specification in Fig. 6 for all children and Fig. 7 for children with immigrant backgrounds. The specification is based on an estimation of Eq. (2), omitting the last not-treated cohort starting school in 2015. The first year of treatment is then a combination of treated and non-treated cohorts from the treatment groups, while the second year of treatment includes the second treated cohort from the 2016 treatment group and the first treated cohort from the 2017 treatment group.

The estimates for test score and proficiency level drop slightly in the first post-treatment year and increase again in the second treatment year, but these changes are not significant. Fig. A.5 in the Appendix presents the event study using the csdid estimator. The pattern is consistent with no apparent effect of treatment in the post-periods. Results for children with an immigrant background are similar, but we see a more jumpy pattern likely reflecting the decrease in sample size.

All in all, our results for child academic outcomes suggest that the free ASP did not impact child cognitive development, neither overall nor in the sub-samples which likely were more affected by the policy. Thus, our results do not lend support to the idea that a

[17] Covid-19 could interfere with our treatment effects since the post-treatment period includes national tests taken in the fall of 2020 and 2021. However, schools were only closed for about 6 weeks, they prioritized children in grades 1–4 when reopening and the number of Covid-19 cases in Norway was relatively low in this period. National test results from 2020 and 2021 show few signs of learning loss neither at the aggregate level nor for particularly vulnerable groups and there is no differential trend for treatment and comparison groups. There are also no signs of differential trends in the share of students taking national tests.

[18] Results using the entire estimation window (2010 and onward) are presented in the Appendix (Table A.2) and are consistent with the main results.

[19] We have run various combinations of treatment and control comparisons, including separately comparing the 2016 treatment group and the 2017 treatment groups to the comparison group, so that early adopters serve as a comparison group for the later adopters. We have also used the 2017 treatment group as a comparison group for the first cohort of the 2016 treatment group. Since the rollout was gradually implemented to less and less socioeconomically disadvantaged city districts, these are potentially the most comparable districts for the first year of treatment. The results are consistent across specifications.

[20] We have also done exploratory analyses of sub-samples by gender and for a sub-sample of children whose mother did not work before the free ASP. We do not find differences by gender, and effects are also not significant for children of non-working mothers.

Table 4
Sub-sample analysis.
Source: Statistics Norway.

	Reading			Mathematics		
	Score	Proficiency	Exempted	Score	Proficiency	Exempted
With immigrant background						
$D^{2016/2017}$	−0.013	−0.026	−0.007	−0.018	−0.012	0.005
	(0.039)	(0.016)	(0.012)	(0.049)	(0.021)	(0.011)
CI with wild bootstrap	[−.100, .0.066]	[−.068, .010]	[−.037, .020]	[−.114, .093]	[−.055, .037]	[−.022, .029]
N	9288	9288	10155	9300	9300	10273
Without immigrant background						
$D^{2016/2017}$	0.030	0.007	0.003	0.025	0.014	−0.000
	(0.047)	(0.010)	(0.005)	(0.033)	(0.010)	(0.004)
CI with wild bootstrap	[−.084, .0.133]	[−.018, .029]	[−.007, .016]	[−.048, .107]	[−.006, .036]	[−.011, .008]
N	23791	23791	24504	23786	23786	24753
Above 60% of median family income						
$D^{2016/2017}$	0.042	0.001	0.006	0.004	0.016+	0.002
	(0.043)	(0.009)	(0.006)	(0.031)	(0.009)	(0.005)
CI with wild bootstrap	[−.067, .0.136]	[−.023, .020]	[−.007, .021]	[−.035, .110]	[−.007, .034]	[−.009, .015]
	24828	24828	25535	24784	24784	25762
Below 60% of median family income						
$D^{2016/2017}$	−0.034	−0.024	−0.017	−0.024	−0.001	−0.003
	(0.049)	(0.019)	(0.010)	(0.042)	(0.023)	(0.008)
CI with wild bootstrap	[−.152, .0.072]	[−.067, .021]	[−.043, .004]	[−.114, .068]	[−.052, .055]	[−.023, .016]
N	8251	8251	9124	8302	8302	9264
Mother is a high school graduate						
$D^{2016/2017}$	0.009	−0.002	0.005	0.022	0.014+	0.001
	(0.043)	(0.010)	(0.005)	(0.029)	(0.008)	(0.003)
CI with wild bootstrap	[−.103, .0.104]	[−.024, .019]	[−.006, .018]	[−.004, .094]	[−.003, .030]	[−.006, .008]
N	25787	25787	26601	25743	25743	26857
Mother is not a high school graduate						
$D^{2016/2017}$	0.030	−0.018	−0.000	−0.007	0.001	0.010
	(0.044)	(0.019)	(0.012)	(0.047)	(0.026)	(0.011)
CI with wild bootstrap	[−.082, .0.122]	[−.061, .026]	[−.029, .025]	[−.111, .107]	[−.058, .061]	[−.014, .034]
N	7292	7292	8058	7343	7343	8169

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Effects for scores are reported as percent of the standard deviation. All panels report results based on Eq. (1). Sub-samples are described in Section 4.3. Standard errors are clustered at the city district level, whereas the CI interval reported is calculated from standard errors using the wild bootstrap procedure. Includes cohorts starting school in Oslo from 2013–2017.

Table 5
Student welfare survey.
Source: Source: Statistics Norway and the Directorate of Education.

	Well-being	Bullying
Two way fixed effects		
$D^{2016/2017}$	−0.023	0.005
	(0.035)	(0.027)
CI w/wild bootstrap	[−.099,.064]	[−.057,.062]
N	29 145	28 871
Callaway St.Anna (csdid)		
ATT	−0.016	−0.030
	(0.042)	(0.029)
CI w/wild bootstrap	[−.097, .0.064]	[−.088, .0.029]
Mean	4.23	4.63
N	29 145	28 871

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Panel 1 reports results based on Eq. (1) and with the outcomes Student well-being and Bullying whereas Panel 2 displays the same outcomes based on a regression using the csdid estimator described in Callaway and Sant'Anna (2021). Standard errors are clustered at the city district level and are estimated using the wild bootstrap procedure (wboot). Confidence intervals are shown in brackets. Includes cohorts starting school in Oslo from 2013–2017.

universal ASP can help reduce educational inequalities, as suggested by Schmitz (2022). However, our results are in line with Felfe and Zierow (2014), who estimate a value-added model and find no effects on school outcomes. They do, however, find positive effects on socio-behavioral development, but this is not possible for us to estimate with the data we have available.

6.2. Exploratory analyses

Although we find no effect of the free ASP on academic performance, the program may have led to other effects for both students and families. Going beyond the pre-analysis plan, we use student welfare surveys to investigate the effect of the rollout on measures of well-being and bullying. We also investigate the effect of the free ASP on maternal labor supply, measured by linear earnings when the child is 6 and 7 years old.

In Table 5, we present results for student well-being and bullying using the two-way fixed effects and csdid estimators (covariates not available). We keep in mind that a high mean value for well-being indicates high well-being and that a high mean value for bullying indicates low bullying. None of the results are significant, although if anything, the signs indicate a somewhat less healthy school socio-emotional environment. Standard errors are substantial, making it hard to draw any clear conclusions.

Table 6 presents results for maternal labor supply using the two-way fixed effects (without and with covariates) and the csdid estimator. Here, we study the overall effect and effects in the sub-samples where we would expect mothers to be more affected, i.e., the immigrant population where we know from other sources that under-representation was substantial before the expansion of the free ASP. We keep in mind that while many children of immigrants were not enrolled in the ASP before the new policy, most other children already attended. For these latter families, the free ASP raised family income compared to the

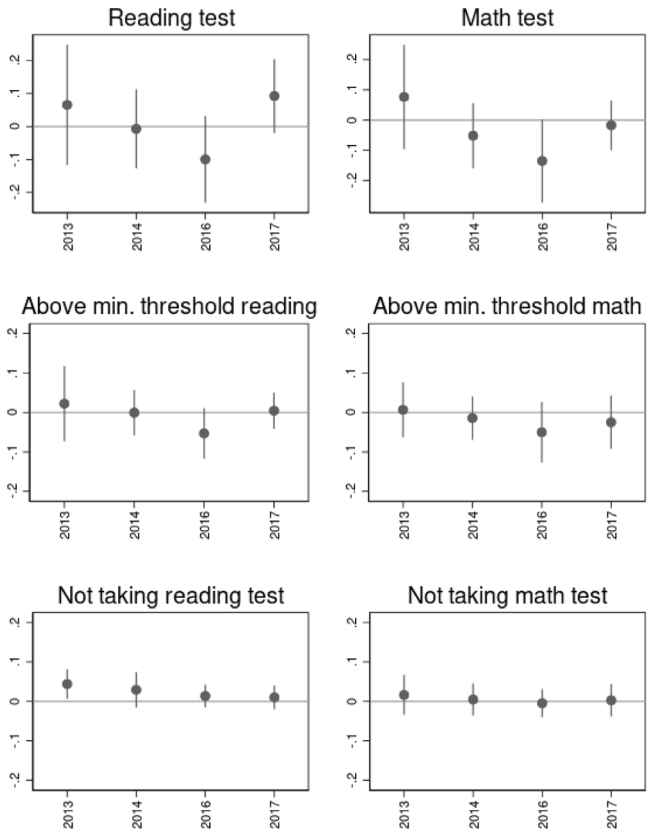


Fig. 7. Event study estimates of student outcomes, children with immigrant background
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

counterfactual with no change in the price of the ASP. We report effects both the year the child turns 6 (and enrolls in school in August) and the year the child turns 7. If there is indeed an effect on mothers' labor supply of the free ASP, we might expect that effects are more pronounced the year the child turns 7, both because we capture 12 months of the free ASP at 7 (compared to August–December the year they turn 6) and because it might take some time to find a job.

In Panel 1, we report the effects on mothers' labor supply for the entire sample. While the direction of the estimate is positive for the two-way fixed effects, it turns negative when we add covariates. The csdid estimator produces a positive estimate, but standard errors are substantial. We next turn to Panel 2 to look at the effects for mothers with immigrant backgrounds. We see that results are positive and quite consistent across specifications. Estimates for the year when the child turns 6 are not significant, but the year the child turns 7 we see that the estimate is larger and precisely estimated when adding covariates. Reassuringly, the csdid estimator produces a similar estimate. The size of the effect is between 1500–2000 USD and constitutes about 10 percent of earnings for this group the year before the new policy. Lastly, in Panel 3, looking at the effects for mothers without an immigrant background, we see a very different pattern. Here, the two-way fixed effects estimates are negative and substantial in both years, suggesting that mothers on the margin reduce their labor supply when the price of the ASP is reduced. However, estimates are imprecise and the csdid estimates do not confirm the larger two-way fixed effects results, and hence we are reluctant to draw strong conclusions for this group.

7. Discussion and conclusion

We study the rollout of a free part-time after-school program in the municipality of Oslo. The take-up was substantial, raising enrollment rates from about 70 to 95% in many affected schools. Using register data that allow us to link children to city districts (and hence treatment status), to their families, and to test score records from national tests in reading and mathematics, we estimate whether the increase in enrollment in ASP affected learning outcomes for students. Our results suggest no overall effects of the program on academic performance. We expected the immigrant population to benefit more from the program, both due to a pre-program lower enrollment rate and as enrollment possibly increased exposure to the Norwegian language. We also expected students from low-income families and students with low-educated mothers to benefit more from the program. However, we find little support for enhanced academic performance for these children. On the flip side, it may be seen as reassuring that results for the population that would have enrolled in the ASP regardless of the free ASP are unchanged. This indicates that the inflow of children from less advantaged backgrounds to the ASP did not generate negative peer effects.

Increased interaction with school peers could still matter for the social environment in school and we therefore use student surveys to investigate the effect of the rollout on well-being and bullying in an exploratory analysis. The results do not lend support to free ASP improving students' socio-emotional welfare in treated schools. Due to data limitations, we cannot carry out the analysis by immigrant background, and hence we are cautious when concluding on this outcome. In a second exploratory analysis, we investigate whether maternal labor supply increases when the family gains access to the free ASP. We find no effect for the overall population, but we show that there is a higher labor supply measured by earnings for immigrant mothers. This suggests that the program, albeit failing to enhance child cognitive development, increased labor supply for a group of mothers who were more likely to enroll their children in the ASP as their children were under-represented before the rollout.

All in all, our results for child academic outcomes suggest that the free ASP did not impact child cognitive development, neither overall nor in the sub-samples which likely were more affected by the policy. Thus, our results do not lend support to the idea that a universal ASP reduces educational inequalities, as suggested by Schmitz (2022). Our results are rather more in line with Felfe and Zierow (2014) who do not find effects on school outcomes. Their study, however, finds a value-added effect on socio-behavioral development. While we do not find support for such effects in our study, the lack of individual data for this outcome and the quality of the measure is a shortcoming.

Why do these studies in somewhat similar contexts find such different results? Previous studies suggest that the quality of ASPs may contribute to explaining positive effects (Kremer et al., 2015). Norwegian ASPs are characterized by free play rather than structured learning activities, which may not be a sufficient tool if the goal is to impact academic results. One striking difference across the contexts seems to be that while the staff/student ratio in our setting is similar to ASPs in Germany, where the most closely related studies of ASPs are from, the share of trained staff is substantially lower (Felfe & Zierow, 2014; Schmitz, 2022). Also, in the German setting, children already in ASPs are often from low-SES backgrounds, and more often immigrants (see e.g., Schmitz, 2022). Thus, these programs may be more suited to engage newly recruited children from such backgrounds. In Norway, there is a positive selection into the ASPs in the first place, and perhaps the program is less suited to enhance the child development of marginalized children, not the least because the educational background of the employees is so low. It may be that more skilled staff could enhance the program quality and accommodate for more structured learning supporting activities in the Norwegian setting. We note, however, that adding learning activities and more structure would imply a more expensive program and lead to increased public expenditure.

Table 6
Maternal labor supply.
Source: Statistics Norway.

	Labor supply, child is 6			Labor supply, child is 7		
	Without controls	With controls	Csdid	Without controls	With controls	Csdid
Panel 1: All						
	0.013	−0.069	0.046	0.037	−0.048	0.099
	(0.067)	(0.052)	(0.078)	(0.076)	(0.054)	(0.087)
CI w/wild bootstrap	[−.121, .0.175]	[−.185, .046]	[−.110, .201]	[−.125, .225]	[−.167, .080]	[−.060, .257]
N	35026	35026	35026	35026	35026	35026
Panel 2: With immigrant background						
	0.093	0.080	0.117	0.169+	0.155*	0.218*
	(0.099)	(0.090)	(0.089)	(0.080)	(0.069)	(0.091)
CI w/wild bootstrap	[−.146, .0.301]	[−.140, .273]	[−.043, .277]	[−.019, .343]	[−.009, .302]	[.057, .378]
N	10273	10273	10273	10273	10273	10273
Panel 3: Without immigrant background						
	−0.058	−0.130+	−0.017	−0.063	−0.138+	−0.015
	(0.067)	(0.066)	(0.082)	(0.070)	(0.065)	(0.069)
CI w/wild bootstrap	[−.215, .0.084]	[−.286, .017]	[−.177, .143]	[−.220, .077]	[−.275, .008]	[−.147, .117]
N	24753	24753	24753	24753	24753	24753

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Results are based on Eq. 1 with the outcome Mothers' linear earnings (adjusted with the basic amount G), without and with covariates respectively, and the same outcome based on a regression using the csdid estimator described in Callaway and Sant'Anna (2021). Standard errors are clustered at the city district level, whereas the CI interval reported is calculated from standard errors using the wild bootstrap procedure. Standard errors using the csdid estimator are estimated with the wild bootstrap procedure (wboot). Includes cohorts starting school in Oslo from 2013–2017.

CRediT authorship contribution statement

Nina Drange: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Astrid Marie Jorde Sandsør:** Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Data availability

The authors do not have permission to share data.

Appendix A

See Figs. A.1–A.5 and Tables A.1–A.4.

Table A.1
Sub-sample analysis with covariates.
Source: Statistics Norway.

	Reading			Mathematics		
	Score	Proficiency	Exempted	Score	Proficiency	Exempted
With immigrant background						
$D^{2016/2017}$	−0.007	−0.025	−0.007	−0.015	−0.012	0.004
	(0.027)	(0.015)	(0.012)	(0.037)	(0.021)	(0.010)
N	9288	9288	10155	9300	9300	10273
Without immigrant background						
$D^{2016/2017}$	0.010	0.003	0.004	0.015	0.011	0.001
	(0.034)	(0.010)	(0.005)	(0.026)	(0.010)	(0.004)
N	23791	23791	24504	23786	23786	24753
Above 60% of median family income						
$D^{2016/2017}$	0.017	−0.005	0.008	0.024	0.012	0.003
	(0.031)	(0.009)	(0.005)	(0.026)	(0.009)	(0.004)
	24828	24828	25535	24784	24784	25762
Below 60% of median family income						
$D^{2016/2017}$	−0.035	−0.026	−0.016	−0.033	−0.007	−0.002
	(0.032)	(0.015)	(0.010)	(0.035)	(0.023)	(0.008)
N	8251	8251	9124	8302	8302	9264
Mother is a high school graduate						
$D^{2016/2017}$	0.010	−0.001	0.005	0.027	0.016+	0.001
	(0.031)	(0.010)	(0.005)	(0.025)	(0.008)	(0.003)
N	25787	25787	26601	25743	25743	26857
Mother is not a high school graduate						
$D^{2016/2017}$	0.031	−0.014	−0.002	0.000	0.004	0.009
	(0.034)	(0.018)	(0.012)	(0.042)	(0.028)	(0.011)
N	7292	7292	8058	7343	7343	8169

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Effects are reported as percent of a standard deviation. All panels report results based on Eq. 1. Covariates are described in Section 4.3. Sub-samples are described in Section 4.3. Standard errors are clustered at the city district level. Includes cohorts starting school in Oslo from 2013–2017.

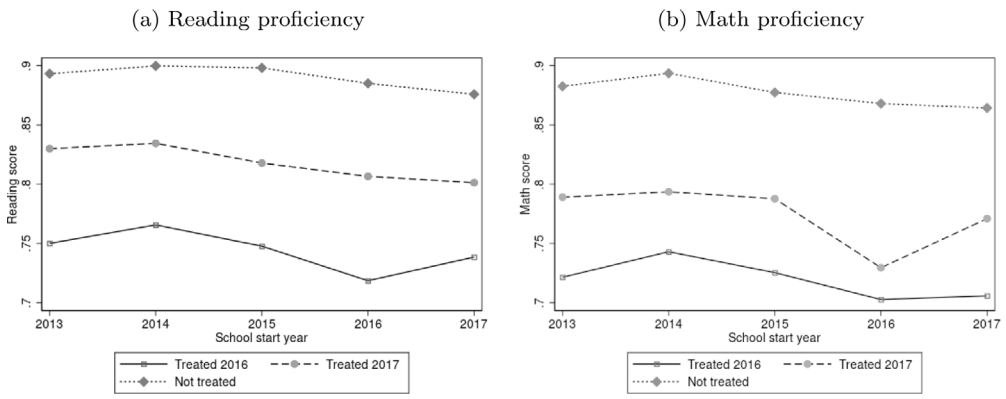


Fig. A.1. Trend in national test proficiency level, all children.
Note: The *x*-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

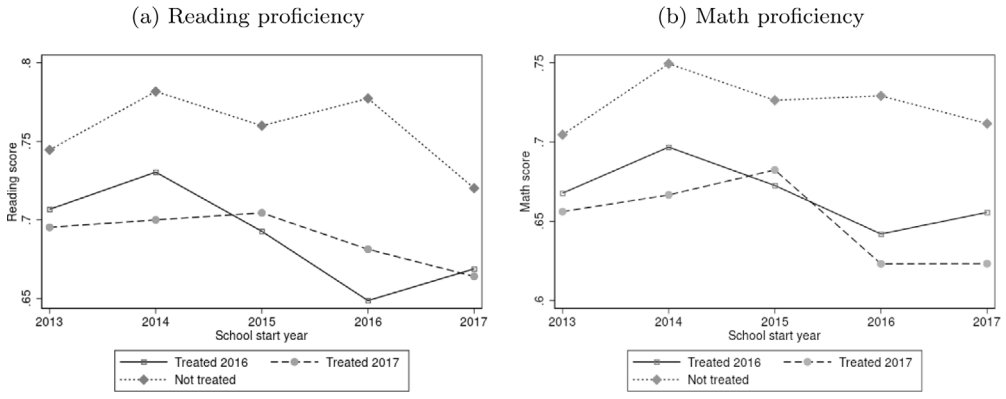


Fig. A.2. Trend in national test proficiency level, immigrant background
Note: The *x*-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

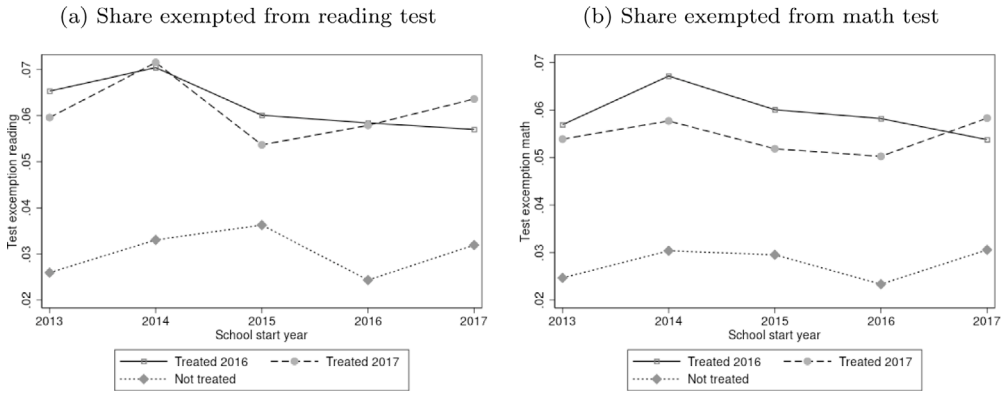


Fig. A.3. Trend in test exemptions, all children
Note: The *x*-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

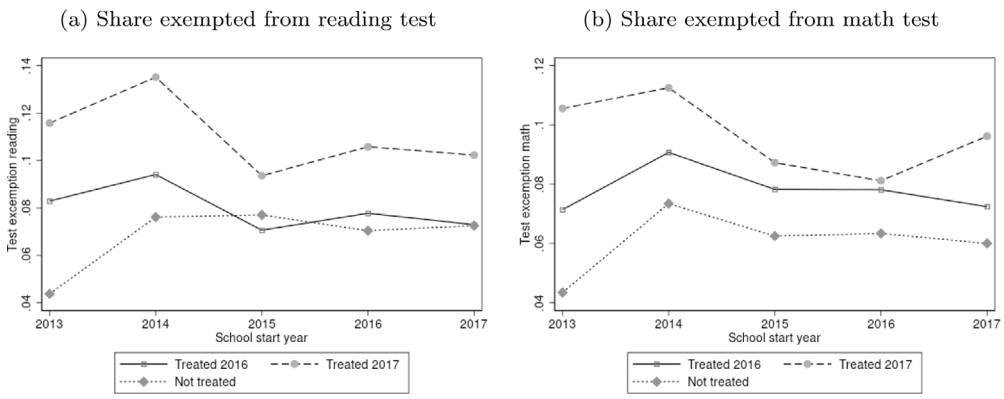


Fig. A.4. Trend in test exemptions, immigrant background
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

Table A.2
Results for estimation period 2010–2017.
Source: Statistics Norway.

	Reading			Mathematics		
	Score	Proficiency	Exempted	Score	Proficiency	Exempted
Panel 1: Two way fixed effects without covariates						
$D^{2016/2017}$	−0.001	−0.017*	−0.008	−0.010	−0.006	−0.007+
	(0.024)	(0.007)	(0.005)	(0.023)	(0.008)	(0.003)
N	51081	51091	53730	51127	51137	54227
Panel 2: Two way fixed effects with covariates						
$D^{2016/2017}$	−0.007	−0.019*	−0.007	−0.010	−0.007	−0.006+
	(0.022)	(0.008)	(0.004)	(0.020)	(0.008)	(0.003)
N	51081	51091	53730	51127	51137	54227
Panel 3: Callaway St. Anna (CSDID)						
ATT	0.053+	−0.000	0.002	0.025	0.012	−0.001
	(0.029)	(0.010)	(0.005)	(0.032)	(0.014)	(0.007)
N	51081	51091	53730	51127	51137	54227

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. Panel 1 reports the results based on Eq. 1, whereas Panel 2 displays outcomes from Eq. 1 and with covariates described in Section 4.3. Panel 3 reports results from the csdid estimator described in Callaway and Sant'Anna (2021). Standard errors are clustered at the city district level. Includes cohorts starting school in Oslo from 2010–2017. In the regressions, the score is standardized, and hence the estimates are reported as percent of a standard deviation.

Table A.3
Results enrollment.
Source: Student Information System (GSI).

	First grade	Second grade	Third grade	Fourth grade
$D^{2016/2017}$	83.7*	96*	138.7*	212.1*
	(24.5)	(25.2)	(26.1)	(36)
Mean pre-intervention	413	283	273	241
N	471	478	483	491

Note: Standard errors in parentheses $+p < 0.10$, * $p < 0.05$. The table reports the results based on Eq. 1. Standard errors are clustered at the city district level. Includes cohorts starting school in Oslo from 2013–2017.

Table A.4
Staff/student ratio before and after the free ASP.
Source: Student Information System (GSI).

	Treated districts		Untreated districts	
	2014		2014	
Trained staff/student ratio	0.019	(0.007)	0.010	(0.004)
Staff/student ratio	0.099	(0.016)	0.103	(0.007)
	2021		2021	
Trained staff/student ratio	0.031	(0.015)	0.022	(0.008)
Staff/student ratio	0.095	(0.015)	0.104	(0.018)

Note: The table shows the trained staff and staff-to-student ratios in 2014 and 2021, separate for treated and untreated city districts. Standard deviations in parentheses.

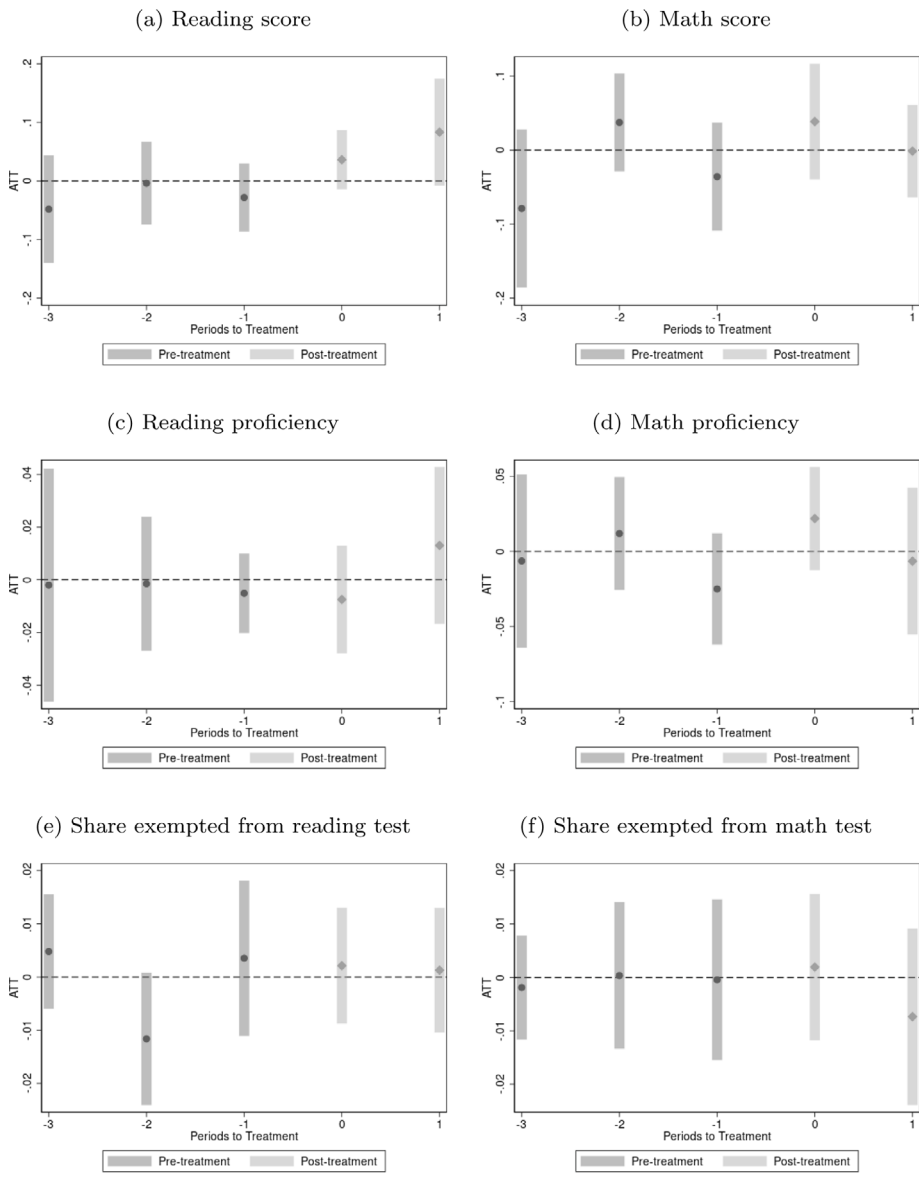


Fig. A.5. Event study estimates csdid, all children
Note: The x-axis corresponds to the year the children start school. The test is taken four years later, i.e., in 2021 for the last observation in the figure.
Source: Statistics Norway.

References

Aizer, A. (2004). Home alone: Supervision after school and child behavior. *Journal of Public Economics*, *88*(9–10), 1835–1848.

Bettinger, E., Hægeland, T., & Rege, M. (2014). Home with mom: the effects of stay-at-home parents on children's long-run educational outcomes. *Journal of Labor Economics*, *32*(3), 443–467.

Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: Who's minding the kids? In *Handbook of the economics of education, vol. 2* (pp. 1163–1278). Elsevier.

Borusyak, K., Jaravel, X., & Spiess, J. (2022). *Revisiting event study designs: Robust and efficient estimation*. CEPR discussion paper DP17247.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Caspersen, J., & Utmo, I. F. (2022). Ny rammeplan i SFO. Kartlegging av innføringen av ny rammeplan i SFO. Delrapport 1. *NTNU Samfunnsforskning*.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, *24*(2), 97–140.

de Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

de Chaisemartin, C., & d'Haultfoeuille, X. (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal*, *26*(3), C1–C30.

Dehos, F. T., & Paul, M. (2023). The effects of after-school programs on maternal employment. *Journal of Human Resources*, *58*(5), 1644–1678.

Drange, N., & Telle, K. (2020). Segregation in a universal child care system: Descriptive findings from Norway. *European Sociological Review*, *36*, 886–901. http://dx.doi.org/10.1093/esr/jcaa026.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, *45*(3), 294–309.

Epple, D., & Romano, R. E. (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics, vol. 1* (pp. 1053–1163). Elsevier.

Felfe, C., Lechner, M., & Thiemann, P. (2016). After-school care and parents' labor supply. *Labour Economics*, *42*, 64–75.

Felfe, C., & Zierow, L. (2014). After-school center-based care and children's development. *The BE Journal of Economic Analysis & Policy*, *14*(4), 1299–1336.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Kautz, T., Heckman, J. J., Diris, R., Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*. NBER working paper 20749.

Kremer, K. P., Maynard, B. R., Polanin, J. R., Vaughn, M. G., & Sarteschi, C. M. (2015). Effects of after-school programs with at-risk youth on attendance and externalizing behaviors: A systematic review and meta-analysis. *Journal of Youth and Adolescence*, *44*(3), 616–636.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, *76*(2), 275–313.

Leuven, E., & Oosterbeek, H. (2018). *Class size and student outcomes in Europe*. EENEE, Analytical Report, (33).

Levine, P. B., & Zimmerman, D. J. (2010). *Targeting investments in children: fighting poverty when resources are limited*. University of Chicago Press.

Martínez, C., & Perticará, M. (2020). Home alone versus after-school programs: the effects of adult supervision on child academic outcomes. *International Journal of Educational Research*, *104*, Article 101601.

Ministry of Children and Families (2012-2013). *Meld. St. 6. (2012–2013). En helhetlig integreringspolitikk. Mangfold og fellesskap*.

Ministry of Education (1998). *Lov om grunnskolen og den vidaregåande opplæringa (opplæringslova)*.

Ministry of Education (2018). *Budsjettspørsmål fra Stortinget. Svar påspørsmål 65 fra Finanskomiteen/SVs fraksjon av 09.10.2018: Gratis SFO*.

Ministry of Education (2019-2020). *Meld. St. 9. (2019–2020). Tett på - tidlig innsats og inkluderende fellesskap i barnehage, skole og SFO*.

Ministry of Education (2020). *Forskrift om rammeplan for skolefritidsordningen*.

Oslo Kommune (2014). *Rammeplan for aktivitetsskolen. Utdanningsetaten, Oslo*.

Oslo Kommune (2018). *Rammeplan for aktivitetsskolen i Oslo. Utdanningsetaten, Oslo*.

PROBA samfunnsanalyse (2017). *Evaluering av forsøk med gratis deltidsplass i AKS*. Rapport 2017-10.

Rios-Avila, F., Sant'Anna, P., & Callaway, B. (2022). *CSDID: Stata module for the estimation of difference-in-difference models with multiple time periods*. Boston College Department of Economics.

Rogne, A. F., Borgen, S. T., & Nordrum, E. (2021). School segregation and native flight: Evidence from School Catchment Area borders. http://dx.doi.org/10.31235/osf.io/xykdg.

Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American Journal of Community Psychology*, *45*(3), 310–324.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the economics of education, vol. 3* (pp. 249–277). Elsevier.

Schanzenbach, D. W. (2020). The economics of class size. In *The economics of education* (pp. 321–331). Elsevier.

Schmitz, L. (2022). *Heterogeneous effects of after-school care on child development*. DIW Berlin Discussion Papers 2006, 55 S..

Seidlitz, A., & Zierow, L. (2020). *The impact of all-day schools on student achievement - evidence from extending school days in German primary schools*. CESifo working paper 8618.

Seppanen, P. S., Love, J. M., deVries, D. K., Bernstein, L., Seligson, M., Marx, F., & Kisker, E. E. (1993). *National study of before- and after-school programs*. Washington, DC: U.S. Department of Education.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Vandell, D. L., Larson, R. W., Mahoney, J. L., & Watts, T. W. (2015). Children's organized activities. In *Handbook of child psychology and developmental science* (pp. 1–40). Wiley Online Library.

Vandell, D. L., Pierce, K. M., & Dadisman, K. (2005). Out-of-school settings as a developmental context for children and youth. In *Advances in child development and behavior, vol. 33* (pp. 43–77). Elsevier.

Varde Hartmark (2018). *Brukerundersøkelse for aktivitetsskoletilbudet i Oslo*.

Wendelborg, C., Caspersen, J., Mordal, S., Ljusberg, A.-L., Valenta, M., & Bunar, N. (2018). *Lek, læring og ikke-pedagogikk for alle: Nasjonal evaluering av skolefritidsordningen i Norge*. NTNU Samfunnsforskning AS.

Zief, S. G., Lauver, S., & Maynard, R. A. (2006). Impacts of after-school programs on student outcomes. *Campbell Systematic Reviews*, *2*(1), 1–51.