CESifo WORKING PAPERS

10215
2023

January 2023

The Effects of a Free Universal After-School Program on Child Academic Outcomes

Nina Drange, Astrid Marie Jorde Sandsør

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10215*

The Effects of a Free Universal After-School Program on Child Academic Outcomes

Abstract

Studies have shown that a lack of adult supervision of school-aged children is associated with antisocial behavior and poor school performance. To mitigate this, one policy response is to provide structured, adult-supervised programs offered after school throughout the academic year. After-school programs in Norway are an integrated part of school, used to extend the school day to a full working day by providing care before and after school. Participation is voluntary and is subject to fees paid by parents. In the past decade, the quality and content of these programs and the role they can play in integrating children have been under scrutiny. In 2016/17, the city 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. We utilize the staggered roll out of this free after-school program to investigate enrollment, learning outcomes and student wellbeing. The take-up was substantial, raising enrollment rates rates from about 70 to 95% in the first wave of affected schools. However, our difference-in-differences estimates show little 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 and we find no evidence of increased maternal labor supply.

JEL-Codes: I210, I240, J130.

Keywords: after-school program, after-school care, difference-in-differences.

*Nina Drange**
Frisch Centre & Statistics Norway
Gaustadalleen 21
Norway – 0349 Oslo
n.e.drange@frisch.uio.no

Astrid Marie Jorde Sandsør
Department of Special Needs Education
University of Oslo / Norway
a.m.j.sandsor@isp.uio.no

*corresponding author

We are grateful to Marte Rønning, Nicolai Topstad Borgen 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), and the Workshop on Educational Resources and Student Performance (Oslo). The preparation of this article was supported by funding from the European Research Council Consolidator Grant ERC-CoG-2018 EQOP (Grant No. 818425) and was part of CREATE - Center for Research on Equality in Education, funded by the Research Council of Norway through its project number 331640.

1 Introduction

Studies have shown that a lack of adult supervision of school-aged children is associated with antisocial behavior (Aizer, 2004) and poor school performance (Bettinger et al., 2014). To mitigate this, one policy response is to provide adult supervision through after-school programs (ASPs). ASPs are structured, adult-supervised programs offered after school throughout the academic year and often during holidays as well. While there is substantial variation in programs across countries, they often supervise and facilitate group-based activities such as homework time, social interaction, snacks, sports and crafts and typically offer more than one activity (Roth et al., 2022; Vandell et al., 2015, 2005).

ASPs are often categorized within the broader category of organized activities (Vandell et al., 2015), or out-of-school time (OST) enrichment activities (Darling-Hammon et al. 2020), including extracurricular activities, sports, community service, drama, summer camps and other school and community-based programs. Studies reviewing the recent literature have shown promising results where attendance is associated with improvements for both academic and socio-emotional student outcomes as well as for family-related outcomes (Vandell et al., 2015; Darling-Hammond et al., 2020; Durlak et al., 2010; Lauer et al., 2006).

Causal evidence on ASPs is, however, scarce. Existing studies suggests that at-risk students benefit from ASPs the most (Levine and Zimmerman, 2010; Schmitz, 2022; Felfe and Zierow, 2014) and that these benefits depend on the quality of the intervention (Kremer et al., 2015). The counterfactual has also been shown to matter: Children who do not have access to adult supervision at home, gain more from ASPs (Martínez and Perticará, 2020). What is still lacking is a greater understanding of whether or how the organization of after-school programs matter and for whom they matter.

After-school programs in Norway are an integrated part of school, used to extend the school day to a full working day by providing care before and after school. 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, Oslo gradually introduced and expanded an offer of free part time participation in its

ASP, starting with city districts with high shares of immigrants. The gradual expansion provides us with a unique opportunity to study how a cost-reducing policy affects ASP enrollment and whether enrollment in turn affects student outcomes. Since the program targeted city districts with initial low enrollment and high shares of children with immigrant background, we can also investigate whether it enhanced skill development for disadvantaged students, potentially narrowing achievement gaps.

We utilize the staggered roll out in a difference-in-differences analysis to investigate whether the introduction of the program led to (1) an increase in enrollment and (2) increased learning. The project was pre-registered at OSF (https://osf.io/qdw9e) prior to obtaining the outcome data. We hypothesized that the effects of the program would be concentrated among those who were most likely to be affected by the introduction of the free ASP: students with an immigrant background and students from low-income families. We expect effects to be driven by enrollment at the extensive margin, though there could also be an effect for children who would have counterfactually enrolled, either though increased enrollment at the intensive margin or reduced costs. Register data allow us to link children to city districts (and hence treatment status) prior to school start, as well as to their families and to test score records from national tests in reading and mathematics taken during the autumn following four (possible) years of ASP. This minimizes attrition and enables a careful analysis of sub-samples by family background.

Going beyond the pre-analysis plan, we use student well-being surveys to investigate the effect of the roll-out 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.

Our findings suggest that the take-up was substantial, raising enrollment rates rates from about 70 to 95% in affected schools. The results, however, suggest little overall effects of the program on academic performance, neither on average nor across subgroups. There is also little evidence of effects on well-being, bullying and maternal labor supply. 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 3 presents the data, Section 5 shows the trends in outcomes and Section 6 presents results. Finally, Section 7 offers 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 1st to 4th 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 or 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 stay in the ASP after school ends until 3:30/4 pm and can attend the program two days 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 been low.

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.

The ASP regulation allows for local variation in organization, content and priority areas and the municipality of Oslo has its own framework for ASP (Oslo Kommune, 2018, 2014). 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 and onward includes four target areas; (1) nature, environment and sustainable development, (2) art culture and creativity, (3) physical activity and (4) food and health. 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 cooperate 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 APSs, leaders of ASPs were in 2022 asked about whether the ASP is an arena where formal language training is carried out. 33% of the leaders respond yes, absolutely, and another 22% respond yes, partly. When asked whether children with minority language backgrounds participate in regular activities, 84% of leaders and 88% of employees respond yes, absolutely, with the remaining almost all answering yes, partly (Caspersen and Utmo, 2022).

In Oslo, parents receive yearly questionnaires about ASPs. The responses from 2018 (Varde Hartmark, 2018) showed that price was a major reason for not participating in APSs. 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 across grade levels, there were noticeable differences across schools on measures of satisfaction, with the same schools typically scoring high/low across categories,

Table 1: The roll-out of the free after-school program across schools

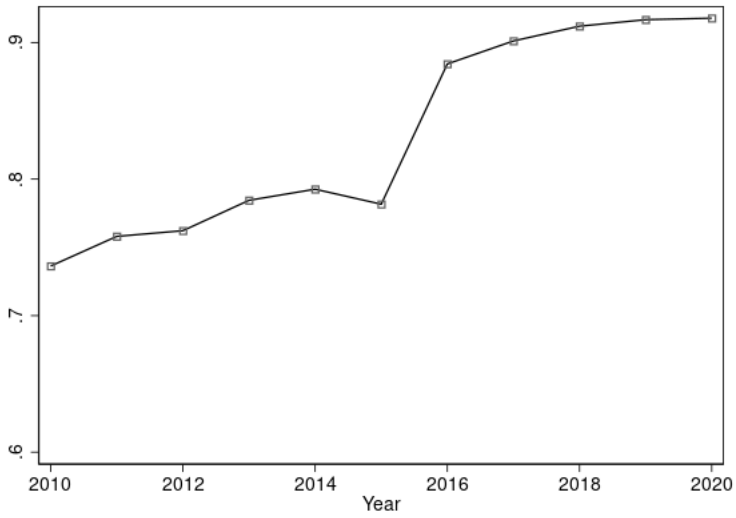
School year	1st grade		2nd grade		3rd grade		4th grade	
	No	Yes	No	Yes	No	Yes	No	Yes
2012/13	102	0	102	0	102	0	102	0
2013/14	100	2	100	2	100	2	100	2
2014/15	99	3	99	3	99	3	99	3
2015/16	99	3	99	3	99	3	99	3
2016/17	68	**34**	102	3	102	3	102	3
2017/18	48	**54**	68	**34**	102	0	102	0
2018/19	37	65	48	**54**	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**

suggesting systematic variation in perceived quality.

The free ASP program. Following initial trial projects in the schools Mortensrud, Tøyen and Vahl, the city of Oslo gradually introduced a free part time ASP, beginning with the most disadvantaged city districts and finally including all of the 15 city districts (102 schools). The roll-out of the program started in the academic year of 2016/2017 in four city districts (34 schools). Children attending 1st grade were eligible the first school year, 1st-2nd in 2017/2018, 1st-3rd in 2018/2019 and 1st-4th in 2019/2020. In 2017/2018, the program was expanded to four new city districts (20 schools), with eligibility expansion 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 grade levels 2 and 3 were eligible for free ASP. The roll-out 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 Figure 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 roll-out, and continued to increase gradually in the following years as the program expanded. In 2016, roll-out increased with about 10 percentage points, from about 80% to 90% and then increased further to about 93% from 2017-2019 before it stabilized.

Figure 1: Share of pupils in AKS - entire Oslo, 1st grade



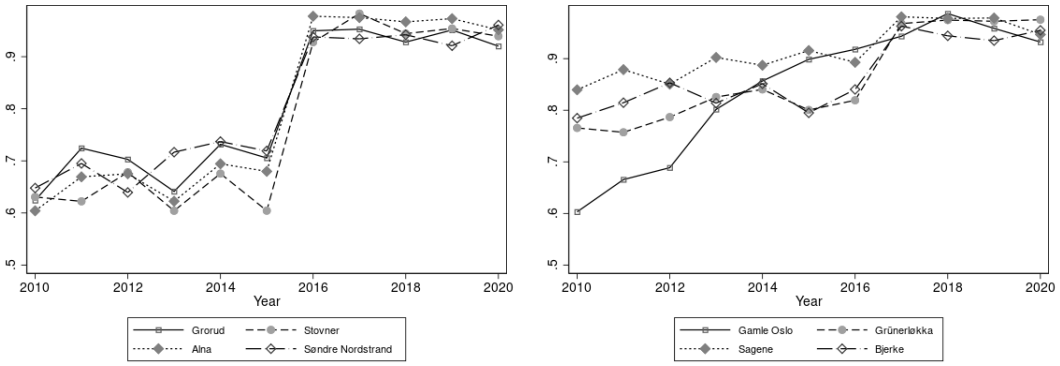
Source: Municipality of Oslo, Department of Education.

Figures 2a and 2b, showing the take-up for the city districts that introduced free ASP in 2016/17 and 2017/18, reflect the same pattern found in Figure 1. In Figure 2a, we see that enrollment increased in particular 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 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 Figure 2b, initial enrollment in ASP was higher and there was also a gradual increase in enrollment occurring from the start of the period, 2010, until the year before the implementation, 2016. Still, we see a jump in enrollment of 5-10 percentage points in 2017 when 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 free ASP for first graders starting in either 2018/19 or 2019/20), initial enrollment was high, already at more than 95% the year prior to introducing free ASP, with little potential for further increase as the program was introduced.

Figure 3 illustrates the take-up rate for grades 1-4 in the same time period, separate for the city districts that introduced the program in 2016 and 2017 and the city districts that introduced

Figure 2: Implementation of free part time after-school programs across city districts

(a) Implemented in 2016/17 (b) Implemented in 2017/2018



Source: Municipality of Oslo, Department of Education and Statistics Norway.

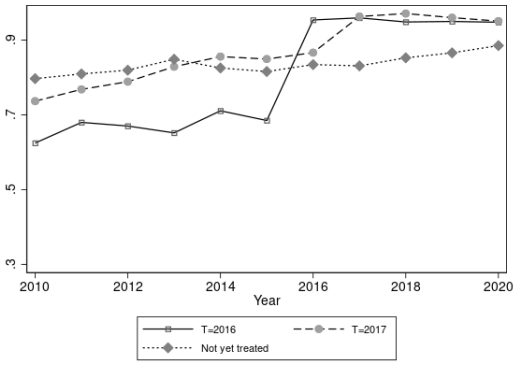
the program later (not yet treated). The figure clearly illustrates the gradual expansion for both treated groups, with increases in enrollment for first grade occurring the first year of the expansion and subsequently, in the second, third and fourth grade for the second, third and fourth year of the expansion. The not yet treated group includes city districts that introduced free ASP for only first graders in 2018/19 or 2019/20 (see Table 1) and for this group we see a slight increase in take-up among first graders at the end of the period. There is also a slight increase for second graders, although this grade was not covered by the expansion.

Together, these figures illustrate that the roll-out 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.

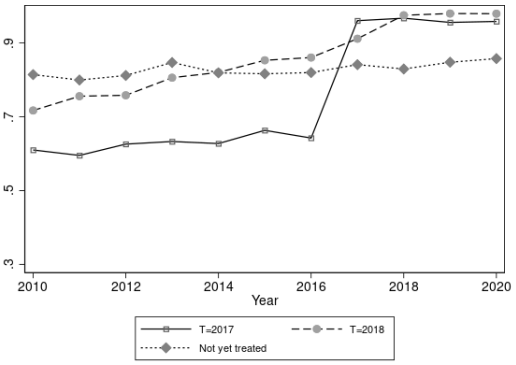
It is important to keep in mind that free part time ASP consists of two different treatments; changing the counterfactual by bringing kids from other forms of care into formalized ASP, to a greater extent occurring in early intervention districts, and reducing the cost for families already using ASP, to a greater extent occurring in later intervention districts. We expect effects to be driven by enrollment at the extensive margin, though there could also be an effect for children who would have counterfactually enrolled, either though increased enrollment at the intensive margin or reduced costs.

Figure 3: Implementation across grades and time

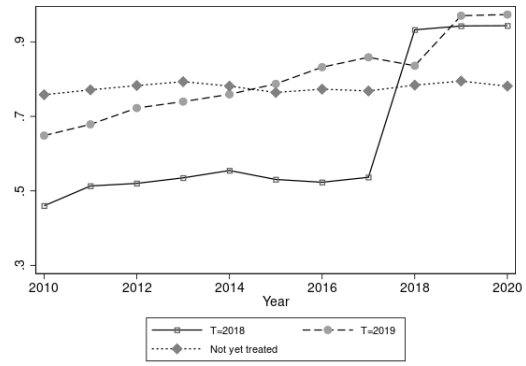
(a) First grade



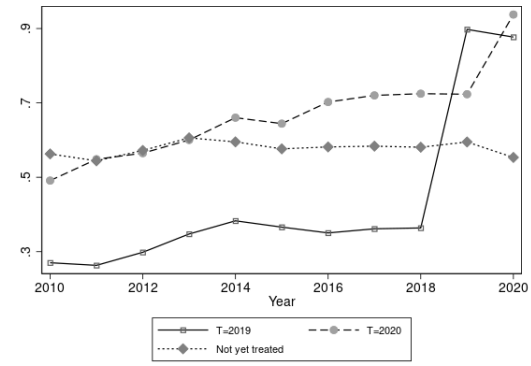
(b) Second grade



(c) Third grade



(d) Fourth grade



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 prior to treatment was not eligible for free ASP at all, whereas a first grader that 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 free part time ASP. This strategy will yield unbiased estimates of the ASP if trends in outcomes of children in treated city districts are similar to trends in comparison districts, if the composition of families stay 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 who introduced free ASP in 2016 and 2017 and school districts that did not introduce free ASP until later.[1]

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

$$(1)\ Y_{i,t} = \alpha_i + \lambda_t + \delta^{DD} D_{i,t} + \eta X_i^{\tau} + \varepsilon_{i,t}$$

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.[2] α_i are city district fixed effects[3] and λ_t are cohort fixed effects. $D_{i,t}$, our variable of interest, is a dummy variable equal to 1 if child i lives in a city district with free ASP the year they start school.[4] 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

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

[2] 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.

[3] 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.

[4] For the case with a single treatment time 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 free ASP during the time period we study and $\lambda_t$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 free ASP was introduced.

father is born abroad, mother's country of origin, parental education and parental 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 city district.

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

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

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 shows how treatment effects develop over time. [5]

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; the first and second treated cohort from the first expansion (test year 2020 and 2021) and 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 2010 and onward.[6] However, after a careful inspection of the trends in the pre-treatment period, our main analysis focus 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 (see discussion below). We report results from analyses where we also include cohorts of school starters 2010-2012 as a robustness check, but trends are more similar in the later cohorts. Lastly, we need to pay particular attention to the three pilot schools where the program was implemented before the roll-out. In the analysis we exclude pupils that take their national test at these schools (the results do not change if we include these schools in our analysis).

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

[5] 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 and Abraham, 2021). For our event study specification, we presenting results using the Callaway & Sant'Anna estimator csdid (Callaway and Sant'Anna, 2021; Rios-Avila et al., 2022), which takes this into account.

[6] 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.

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 by presenting results using the Callaway & Sant'Anna estimator csdid (Callaway and 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 free part-time ASP in the first wave of the roll-out.

In our pre-registration, we only included academic outcomes. However, we have added an exploratory analysis where we explore whether the expansion of ASPs affected mothers labor supply. We estimate Equation 1 for two separate outcomes: mothers' income 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 on a yearly basis by the Directorate of Education. Fifth grade students answer questions about their experience of being in school and the school social environment. We study the two topics covered in the survey that we think holds relevance for our purpose, bullying and well-being, and estimate effects using Equation 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 at the 1st of January in 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 parental income and educational attainment, as well as mother's continent of origin.

The roll-out of the program is linked by school identifiers and year. The data on the roll out of the free program was collected from the municipality of Oslo and includes information on 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 had a trial project where they got access to free

ASP prior to the roll-out. In these schools there were also an emphasize on the content of the program and hence we focus our analyses on a sample where children taking tests in these schools are excluded.[7]

To study effects of the free after-school program on student welfare, we use data from a survey called Elevundersøkelsen (the Student Survey) that takes place on a yearly basis in the autumn of 5th grade.[8] We have individual data linked to school identifiers, but these are anonymized and cannot be linked to the rest of the data. We can hence not study 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 2013-2017.

4.2 Outcome Variables

The main outcome variables are collected from compulsory national tests in the subject reading and mathematics. These tests are taken at the start of 5th grade, i.e. the cohort starting school in 2013 is 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 psychometric and are designed to capture the full range of skills in these subjects. The results are mainly used to collect information about students' basic 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 exempt. Note that the registry data allows us to follow children that 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 prior to school start (and is registered in our sample), may move to another part of the country, but still be registered with academic outcomes in our data.

[7] Results change very little if these schools are included.

[8] The Student Survey can be carried out in all grades from 5th and onward. It is compulsory for the schools to carry it out in 7th, 10th and 11th grade, whereas it is decided on the school level whether to be carried out in other grades. 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 are very similar and the reduction in sample size is mainly caused by non-response among students. As we are not able to link this data to register data, we are unfortunately not able to investigate the characteristics of non-responders.

We show separate estimates for reading and mathematics using both the scaled score as well as a dummy variables for each subject that measures if the individual scores 2 or 3 (and not 1). In addition to the test scores, we construct dummy variables that captures if the child was exempted from the test.

To explore whether the free ASP program induced some mothers to work more, we construct a linear earnings measure capturing mothers' linear earnings[9] the year their child(ren) in the sample turns 6 and 7 respectively. We note that due to the not pre-registered status of this outcome, this analysis should be regarded as exploratory.

In another exploratory analysis, we look at potential effects on student welfare and perceived bullying. The welfare measure is based on one single question: "Do you thrive in school?"[10] and the answer is graded from 5 (Yes, absolutely), to 1 (Not at all). The bullying outcome is based on the rowmean 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?".[11] 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 are measured the year before the child start 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 birth quarter. Immigrant background is defined as having two parents born abroad. We also construct dummies for 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.[12]

[9] We adjust income by the basic amount (Grunnbeløpet), which is adjusted for inflation and wage growth by the Norwegian parliament every year and is included in most formulae for public welfare transfers.

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

[11] 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?"

[12] Due to a 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 2: Summary statistics

	Treated 2016	Treated 2017	Comparison
Girl	0.49	0.50	0.49
Immigrant	0.54	0.36	0.14
Continent			
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
M finished high school	0.57	0.72	0.89
F finished high school	0.56	0.67	0.86
M university	0.34	0.56	0.75
F university	0.30	0.47	0.68
M edu unknown	0.09	0.07	0.03
F edu unknown	0.10	0.10	0.05

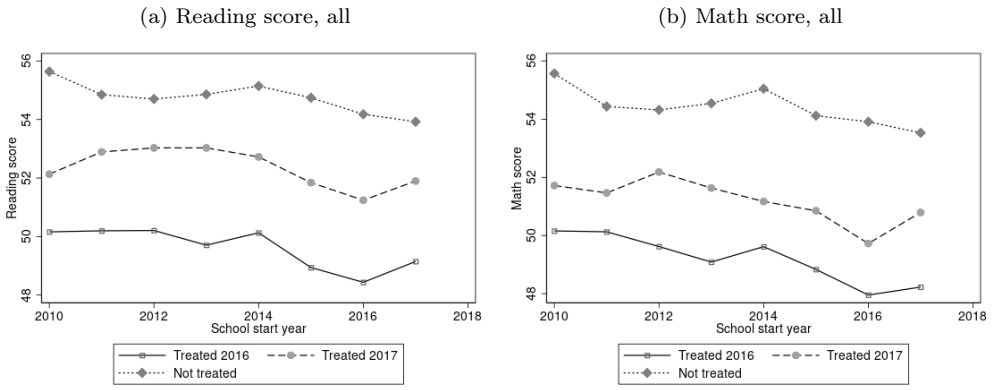
Note: The first column reports background for children residing in city districts that got access to the free ASP in 2016, the second displays background for children in city districts that got access in 2017 whereas the third column displays background for those residing in city districts that got access later. Construction of covariates is described in Section 4.3. Source: Statistics Norway.

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 roll-out started with the city districts with the lowest socioeconomic backgrounds and continued on to the next in line. The identifying strategy when using difference-in-differences, however, hinges on common trends rather than common levels as well as 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. Figures 4a and 4b show the average national scaled test score for reading and mathematics, respectively, for cohorts starting first grade from 2010 to 2017, separately for the 2016 and 2017 treatment groups and the comparison group. Although scores in reading and

Figure 4: Trend in National Test Scores



Source: Statistics Norway.

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 prior to treatment, 2013-2015, while the trends are less stable prior to 2013.

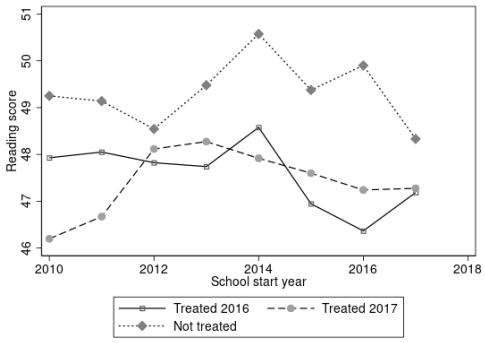
For the 2016 treatment group, there are two cohorts in the post-treatment period. For both reading and mathematics, test scores in the first cohort (started school in 2016) relative to the comparison group do not seem to change, but there is a slight relative increase for the second cohort (started school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

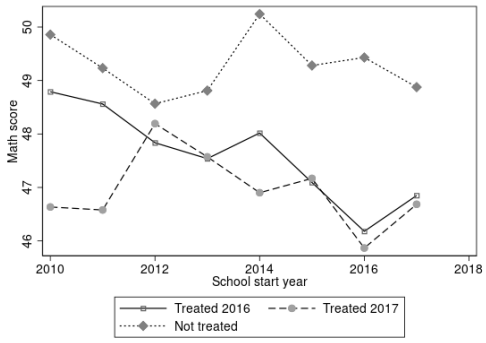
However, average effects could conceal important heterogeneity. Specifically, our hypothesis was that children with immigrant background would benefit the most from the policy, both due to their low relative enrollment prior to the introduction of free ASP and because enrollment would increase their 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 Figures 5a and 5b we therefore look at the same trends for the immigrant population only. The

Figure 5: Trend in National Test Scores, Immigrants

(a) Reading score, immigrant background



(b) Math score, immigrant background



Source: Statistics Norway.

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 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.

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 Figures A.1a and A.1b for the whole population and Figures A.2a and A.2b for immigrants 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 free ASP on the first cohort of treated first graders (2016) while there is a slight relative increase in the second year where both treatment groups (2017) have treated first-fourth graders.

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 potentially could 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 belongs to the lower end of the distribution. We hence go on to investigate trends for whether students are exempt from the test (see Figures A.3a and A.3b for the whole population and Figures A.4a and A.4b for immigrants in the Appendix). Here the trends are less stable in the pre-treatment period. In the post-treatment period, there is 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.

6 Results

6.1 Main results

In Table 3, first and second panel, we present results based on an estimation of Equation (1), with and without covariates. Outcome variables are scaled scores, normalized to have mean 0 and standard deviation 1 for the entire time 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.[13] In light of the new developments in the difference-in-differences literature, we go on to present results from the Callaway & Sant'Anna estimator csdid in Panel 3.[14] The results reflect what we observed when inspecting the trend figures - for most outcomes there is no measurable effect of introducing 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. In Panel 3, where we show results from the specification that implement the csdid estimator, all estimates are either zero or have a positive sign. When we run the csdid estimator, we see a positive effect of 0,053 standard deviations for reading, significant at the 10% level, whereas the other estimates are mostly small and not significant.

As mentioned previously, average effects could conceal important heterogeneity. We are

[13] Results using the entire estimation window (2010/11 and onward) are presented in the appendix (Table A.2) and are consistent with the main results.

[14] We have also 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 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 for the roll-out was gradually implemented to less and less socioeconomically disadvantaged city districts, these are potentially the most comparable districts for the first year of treatment. Results are consistent across specifications.

Table 3: Main results

	Reading			Mathematics		
	Score	**Proficiency**	**Exempted**	**Score**	**Proficiency**	**Exempted**
Panel 1: Two way fixed effects						
$D^{2016/2017}$	0.014	-0.009	-0.002	0.014	0.004	-0.001
	(0.030)	(0.008)	(0.005)	(0.023)	(0.008)	(0.004)
N	33079	33079	34659	33086	33086	35026
Panel 2: Two way fixed effects with covariates						
$D^{2016/2017}$	0.005	-0.013	-0.001	0.008	0.002	-0.000
	(0.027)	(0.008)	(0.005)	(0.021)	(0.009)	(0.004)
N	33079	33079	34659	33086	33086	35026
Panel 3: Callaway & Sant'Anna (csdid)						
ATT	0.053+	0.000	0.002	0.025	0.012	-0.001
	(0.029)	(0.01)	(0.005)	(0.032)	(0.014)	(0.007)
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 results based on Eq. 1, Panel 2 displays outcomes from Eq. 1 and with covariates described in Section 4.3, whereas Panel 3 reports results from the csdid estimator described in Callaway and Sant'Anna (2021). Includes cohorts starting school in Oslo from 2013-2017. Mean refers to children in city districts that receives treatment in 2016 or 2017, in 2015. Source: Statistics Norway.

particularly interested in whether children with immigrant background benefit from the policy, in addition to children from families with low income or where the mother is a high school drop-out. Also, 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 the effect is not necessarily clear (Epple and Romano, 2011; Schanzenbach, 2020; Leuven and Oosterbeek, 2018; Sacerdote, 2011).

Results using the two way fixed effects estimator without covariates are presented in Table 4 and show few signs of heterogeneous effects. 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 for the inclusion of covariates. Note that while the lack of effects suggest 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 already enrolled is unchanged. This indicates that the inflow of kids from less advantaged backgrounds to the ASPs did not have a negative spill-over on the kids that were already enrolled.

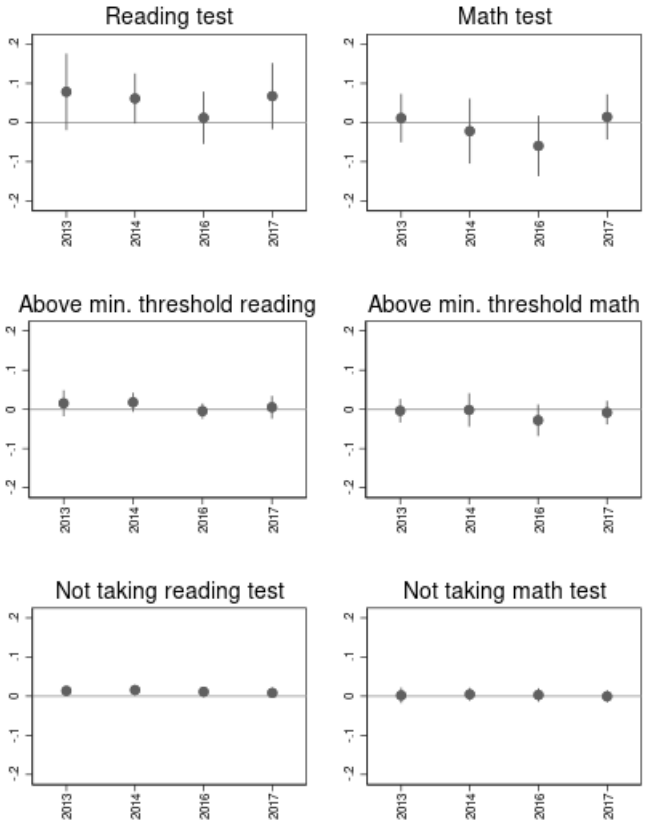
Next, we present results for the event study specification in Figure 6 for all children and Figure 7 for children with immigrant backgrounds. The specification is based on estimation

Table 4: Sub-sample analysis

	Reading			Mathematics		
	Score	**Proficiency**	**Exempted**	**Score**	**Proficiency**	**Exempted**
With immigrant background						
$D^{2016/2017}$	-0.010	-0.026	-0.007	-0.015	-0.012	0.005
	(0.031)	(0.016)	(0.012)	(0.041)	(0.021)	(0.011)
N	9288	9288	10155	9300	9300	10273
Without immigrant background						
$D^{2016/2017}$	0.024	0.007	0.003	0.020	0.014	-0.000
	(0.037)	(0.010)	(0.005)	(0.027)	(0.010)	(0.004)
N	23791	23791	24504	23786	23786	24753
Above 60% of median family income						
$D^{2016/2017}$	0.033	-0.000	0.007	0.033	0.015	0.002
	(0.034)	(0.009)	(0.006)	(0.026)	(0.009)	(0.005)
	24828	24828	25535	24784	24784	25762
Below 60% of median family income						
$D^{2016/2017}$	-0.027	-0.022	-0.017	-0.021	-0.001	-0.003
	(0.037)	(0.018)	(0.010)	(0.034)	(0.023)	(0.008)
N	8251	8251	9124	8302	8302	9264
Mother is a high school graduate						
$D^{2016/2017}$	0.007	-0.002	0.005	0.018	0.014+	0.001
	(0.034)	(0.010)	(0.005)	(0.024)	(0.008)	(0.003)
N	25787	25787	26601	25743	25743	26857
Mother is not a high school graduate						
$D^{2016/2017}$	0.024	-0.018	-0.000	-0.006	0.001	0.010
	(0.035)	(0.019)	(0.012)	(0.039)	(0.026)	(0.011)
N	7292	7292	8058	7343	7343	8169

Note: Standard errors in parentheses + p<0.10, * p<0.05. All panels report results based on Eq. 1. Sub-samples are described in Section 4.3. Includes cohorts starting school in Oslo from 2013-2017. Source: Statistics Norway.

Figure 6: Event study estimates of student outcomes, all children



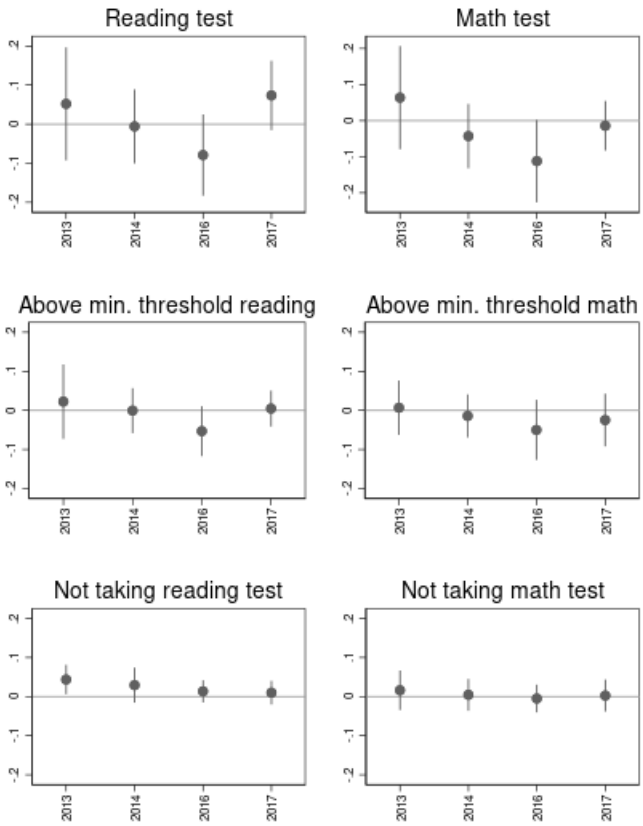
Source: Statistics Norway.

of Equation 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 slightly drop in the first post-treatment year and increase again in the second treatment year, but these changes are not significant. Figure 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.

Figure 7: Event study estimates of student outcomes, children with immigrant background



Source: Statistics Norway.

Table 5: Student well-being survey

	Well-being	Bullying
Two way fixed effects		
$D^{2016/2017}$	-0.023	0.005
	(0.035)	(0.027)
N	29145	28871
Callaway St.Anna (csdid)		
ATT	-0.016	-0.030
	(0.042)	(0.029)
Mean	4.23	4.63
N	29145	28871

Note: Standard errors in parentheses + p<0.10, * p<0.05. Panel 1 reports results based on Eq. 1 and with the outcome 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). Includes cohorts starting school in Oslo from 2013-2017. Source: Statistics Norway and the Directorate of Education.

6.2 Exploratory analyses

Although we find no effect of 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 well-being surveys to investigate the effect of the roll-out on measures of well-being and bullying. We also investigate the effect of the roll-out 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). None of the results are significant, although if anyting, 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 mother's labor supply using the two way fixed effects (without and with covariates) and csdid estimators. Again, the estimates are not significant. While the direction of the estimate is positive for the two way fixed effects and csdid estimator, adding controls changes the direction of the estimate. These analyses suggest that, to the extent that our data allow us to investigate other outcomes, there does not seem to be an effect of the roll-out of free ASP on socio-emotional measures for students or the labor supply of mothers.

7 Conclusion

We study the roll-out of a free after-school program in the municipality of Oslo. The take-up was substantial, raising enrollment rates rates from about 70 to 95% in many affected schools. Using

Table 6: Mothers' labor supply

	Labor supply, child is 6		Labor supply, child is 7	
	Without covariates	**With covariates**	**Without covariates**	**With covariates**
Panel 1: Two way fixed effects				
$D^{2016/2017}$	0.013	-0.069	0.037	-0.048
	(0.067)	(0.052)	(0.076)	(0.054)
N	35026	35026	35026	35026
Panel 2: Callaway St.Anna (csdid)				
ATT	0.046		0.099	
	(0.081)		(0.086)	
Mean	3.42		3.49	
N	35026		35026	

Note: Standard errors in parentheses + p<0.10, * p<0.05. Panel 1 reports results based on Eq. 1 with the outcome Mothers' linear earnings (adjusted with the basic , without and with covariates respectively. Panel 2 displays the same outcome based on a regression using the csdid estimator described in Callaway and Sant'Anna (2021). Includes cohorts starting school in Oslo from 2013-2017. Source: Statistics Norway.

registry 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 non-working mothers to benefit more from the program. However, we find little support for enhanced academic performance for such children. On the flip-side, it may be seen as reassuring that results for the population already enrolled in ASPs is unchanged. This indicates that the inflow of kids from less advantaged backgrounds to the ASPs did not have a negative spill-over on the kids that were already enrolled.

Previous studies suggest that the quality of ASPs may contribute to explain positive effects (Kremer et al., 2015). Norwegian ASPs are characterized by free play rather than structured learning activities, which may not be sufficient tools if the goal is to impact academic results. This program may not have sufficient quality for us to expect enhanced language proficiency. We note, however, that adding learning activities and more structure would likely imply a more expensive program and lead to increased public expenditure.

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 roll-out on well-being and bullying. The results do not support that the ASP improved the socio-emotional well-being in the schools. We also investigated whether mothers' labor supply increased when gaining access to the free ASP, but find no evidence to support this.

The free part-time ASP did not improve child outcomes, but the program did succeed in enrolling new children. The increased presence of children from disadvantaged backgrounds may have potential for further skill development if the program quality is improved through more structured learning supporting activities or small group instruction. Still, it is unclear whether such potential improvement would be sufficient to defend the needed increase in public expenditure.

References

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

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

Borusyak, K., X. Jaravel, and J. Spiess (2021): "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*.

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

Caspersen, J., and I. F. Utmo (2022): "Ny rammeplan i SFO. Kartlegging av innf/oringen av ny rammeplan i SFO. Delrapport 1.,"Technical report, NTNU Samfunnsforskning.

de Chaisemartin, C., and X. D'Haultfœuille (2022): "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey," *NBER Working Paper*.

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

De Chaisemartin, C., and X. d'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–96.

Durlak, J. A., R. P. Weissberg, and M. Pachan (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, 294–309.

Epple, D., and R. E. Romano (2011): "Peer effects in education: A survey of the theory and evidence," in *Handbook of Social Economics* Volume 1: Elsevier, 1053–1163.

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

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

Kremer, K. P., B. R. Maynard, J. R. Polanin, M. G. Vaughn, and C. M. Sarteschi (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, 616–636.

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

Leuven, E., and H. Oosterbeek (2018): "Class size and student outcomes in Europe," *EENEE, Analyt ischer Bericht, (33).*

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

Martínez, C., and M. Perticará (2020): "Home alone versus after-school programs: the effects of adult supervision on child academic outcomes," *International Journal of Educational Research*, 104, 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)*, https://lovdata.no/dokument/NL/lov/1998-07-17-61/KAPITTEL$_1$5

——— (2019-2020): *Meld. St. 9. (2019-2020). Tett på - tidlig innsats og inkluderende fellesskap i barnehage, skole og SFO.*

——— (2020): *Forskrift om rammeplan for skolefritidsordningen.*, https://lovdata.no/dokument/SF/forskrift/2021-05-31-1825.

Oslo Kommune (2014): *Rammeplan for Aktivitetsskolen. Utdanningsetaten, Oslo.*, http://www.svendstuen-sfo.no/wp-content/uploads/rammaplan-AKS.pdf.

——— (2018): *Rammeplan for Aktivitetsskolen i Oslo. Utdanningsetaten, Oslo.*, https://rustad.osloskolen.no/siteassets/aks/2018-rammeplan-for-aktivtetsskolen-i-oslo.pdf.

Rios-Avila, F., P. Sant'Anna, and B. Callaway (2022): "CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods."

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2022): "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *arXiv preprint arXiv:2201.01194*.

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* Volume 3: Elsevier, 249–277.

Schanzenbach, D. W. (2020): "The economics of class size," in *The Economics of Education*: Elsevier, 321–331.

Schmitz, L. (2022): "Heterogeneous effects of after-school care on child development," *DIW Berlin Discussion Paper No. 2006*.

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

Vandell, D. L., R. W. Larson, J. L. Mahoney, and T. W. Watts (2015): "Children's Organized Activities," *Handbook of Child Psychology and Developmental Science*, 1–40.

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

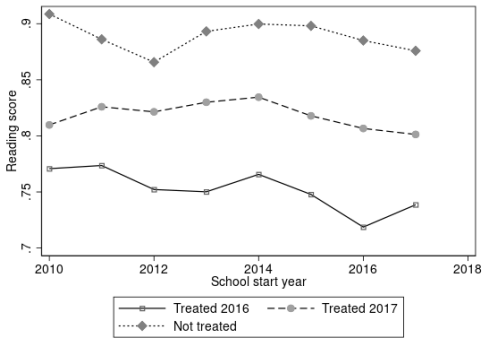
Varde Hartmark (2018): *Brukerundersøkelse for aktivitetsskoletilbudet i Oslo.*

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

Appendix

Figure A.1: Trend in National Test Proficiency level, all

(a) Reading proficiency, all children



(b) Math proficiency, all children

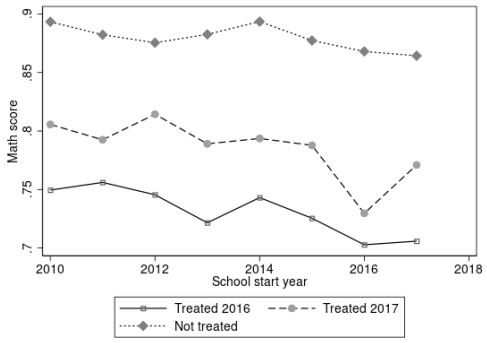
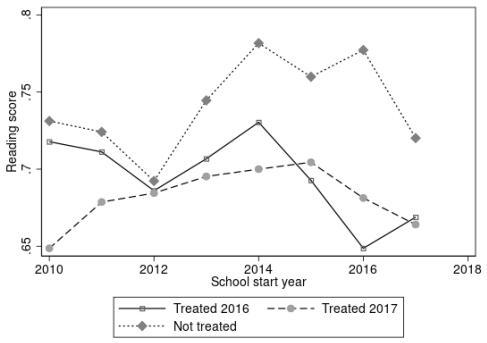


Figure A.2: Trend in National Test Proficiency level, Immigrant background

(a) Reading proficiency, immigrant background



(b) Math proficiency, immigrant background

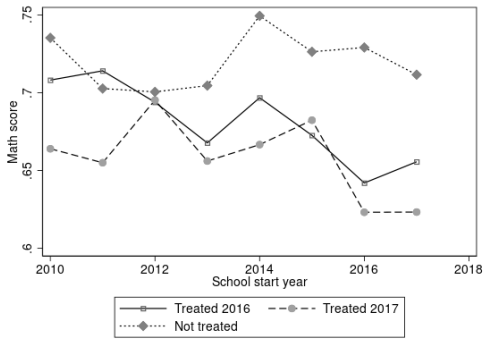
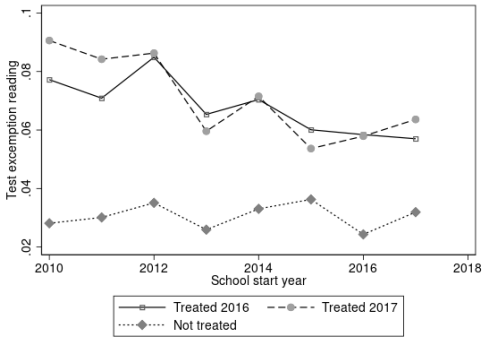
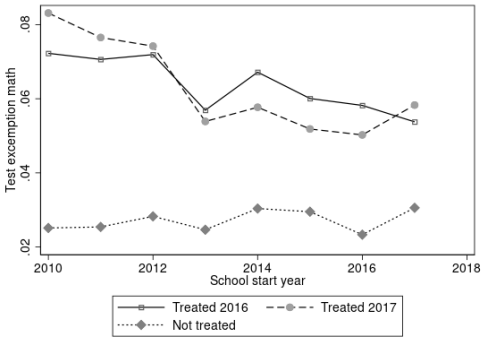


Figure A.3: Trend in test exemptions, all children

(a) Share exempted from reading test



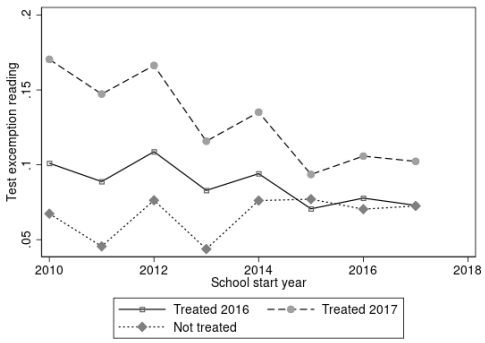
(b) Share exempted from math test



0treatxy201

Figure A.4: Trend in test exemptions, immigrant background

(a) Share exempted from reading test



(b) Share exempted from math test

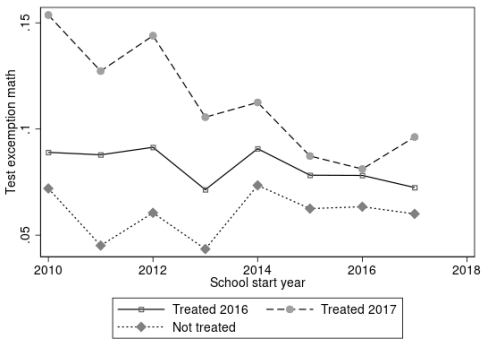


Table A.1: Sub-sample analysis with covariates

	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. Panels report results based on Eq. 1 and sub-samples and covariates are defined in Section 4.3 Source: Statistics Norway.

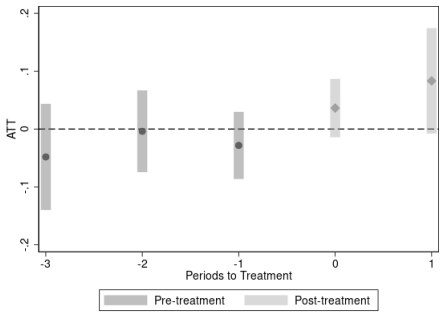
Table A.2: Results - Estimation period is school start year 2010-2017

	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

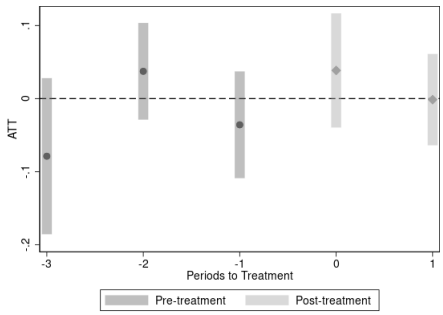
Standard errors in parentheses + p<0.10, * p<0.05. Panel 1 reports results based on Eq. 1, Panel 2 displays outcomes from Eq. 1 and with covariates described in Section 4.3, whereas Panel 3 reports results from the csdid estimator described in Callaway and Sant'Anna (2021). Includes cohorts starting school in Oslo from 2010-2017. Source: Statistics Norway.

Figure A.5: Event study estimates CSDID, all children

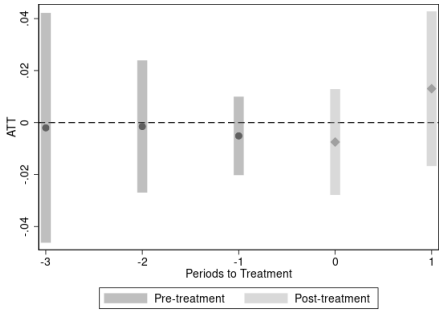
(a) Reading



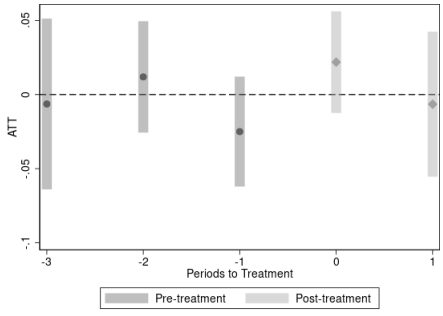
(b) Math



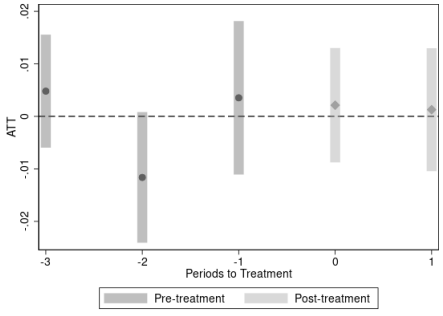
(c) Reading proficiency



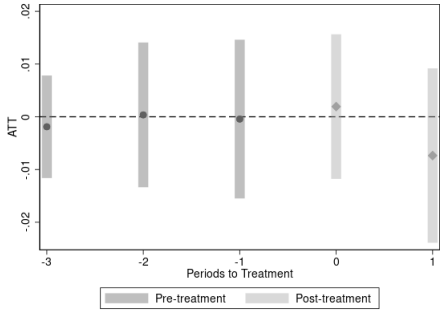
(d) Math proficiency



(e) Excempted Reading



(f) Excempted Math



Source: Statistics Norway.