# The effects of a free universal after school program on 

# child academic outcomes 

Nina Drange, Marte Rønning and Astrid M. J. Sandsør


#### 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 adult supervision through 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, increasing attention has been paid to the quality and content of these programs and the role they can play in integrating children, particularly from immigrant backgrounds. As a result, the city of Oslo has gradually introduced and expanded an offer of free part time participation in this program, starting with city districts with high immigrant shares. We utilize the staggered roll out of this free after-school program to investigate attendance and learning outcomes for students. Our difference-in-difference estimates suggest that the takeup was substantial, raising enrollment rates rates from about 70 to $95 \%$ in affected schools. Preliminary results suggest little overall effect of the program on academic performance.


Keywords: After-school care, difference-in-differences

## 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, as well as during holidays. While there is substantial variation in programs across countries, they often supervise and facilitate activities such as homework time, social interaction, snacks, sports and crafts.

Evidence on ASPs is scarce, but 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 matters: Children who do not have access to adult supervision at home, gain more from ASPs (Martínez and Perticará, 2020).

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, increasing attention has been paid to the quality and content of these programs and the role they can play in integrating children, particularly from immigrant backgrounds. As a result, Oslo has gradually introduced and expanded an offer of free part time participation in this program, starting with city districts with high immigrant shares. We utilize the staggered roll out to investigate whether the introduction of the free program (1) led to an increase in attendance and (2) led to increased learning outcomes for students. Unique registry data 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 taken during the autumn following four (possible) years of ASP. This minimizes attrition and enables a careful analysis of sub-samples after family background. We expect that the effects of the program will be concentrated among those who were most likely to be affected by the introduction of the free after school program, including students from low income families, students with immigrant backgrounds and students with non-working mothers. The study is pre-registered at OSF (https://osf.io/qdw9e).

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, both on average and across subgroups.

## Institutional Context and Reform Details

In Norway, 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, most often taking place on school grounds. The programs are organized at the municipality level, and may be run by both private or public providers. This results in varying costs and content across municipalities, and to some degree also across schools within a municipality. After school programs are viewed as an important arena for aquiring social skills and and for enhance language development, particularly among children who speak another language than Norwegian at home. Most children will attend the ASP, particularly during their first school years, but the enrollment rate for children from immigrant families has been low.

With a full time slot, a child can ASPs before school, at 7:30 am, and stay from the end of the school day until 5 pm . In addition, the child can attend the program during all school holidays, except in July when school is closed for the summer. A half time time slot implies that the child can stay in the after school program from the end of the school day until 4 pm . In addition, the child can attend the program two days during school holidays. Payment depends on family income as 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 20000 and 35000 EURO and 60 EURO for income below 20000 EURO. The fee for a half time slot was $50 \%$ of a full time slot.

Following an initial trial project including three schools in the school years 2013/14-2015/16, the city of Oslo gradually expanded the free half time after school program (free ASP), starting with the most disadvantaged city districts and finally including all of the 15 city districs (102 schools). The roll-out of the program started in the academic year of 2016/2017 in four city districs ( 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 and onwards. In 2017/2018 the program was expanded to four new city districts ( 20 schools), with eligibility expansion following

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

| School year | 1st grade |  |  | 2nd grade |  |  |  | 3rd grade |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |

the first grade cohort as they moved through the school system in the same way as the first expansion. In 2018/2019 two more school districts were included in the program (11 schools), but without the same exansion 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. 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 attendance. As illustrated in Figure 1, enrollement 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 followin years as the program expanded. In 2016, roll-out increased with about 10 percentage points, from $84 \%$ to $94 \%$, and then increased further to about $93 \%$ from $2017-2019$ before dropping somewhat in 2020, likely due to the pandemic. Figures 2a-2c for the city districts that introduced free ASP in 2016/17, 2017/18 and 2018/19 respectively, reflect the same pattern found in Figure 1. In Figure 2a, we see that enrollement increased in particular in the four city districs 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 where entrollement 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 is higher, and there is a gradual increase in enrollment occuring from the start of the period, 2009, until the year before the implementation, 2016. Still, we see a jump in enrollment of 5-10 percentage

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

points in 2017 when free ASP was introduced, with enrollement rates stabalizing around $95 \%$ thereafter. Figure 2c shows enrollement rates for the two city districts that introduced free ASP in 2018/19; St. Hanshaugen and Frogner. Initial enrollement was even higher in these city districts, already at more than $95 \%$ the year prior to introducing free ASP, with little potential for further increase as the program was introduced. The remaining city districts introduced free ASP for first graders in 2019/20.

Together, these figures illustrate that the roll-out was implemented such that city districts with lower enrollement rates initially, typically city districts with a larger share of low income families and immigrants, were treated first. Also, the policy was very effective at increasing enrollment in these early districs relative to later districts where enrollement was already high. It is important to keep in mind that free half 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 occuring in early intervention districts, and reducing the cost for families already using ASP, to a greater extent occuring in later intervention districs.

Figure 2: Implementation of free half time after school programs across city districs


Method

The gradual expansion of 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 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) $Y_{i, t}=\alpha_{i}+\lambda_{t}+\delta^{D D} D_{i, t}+\eta X_{i}^{\tau}+\varepsilon_{i, t}$
where $Y_{i, t}$ is the result from national tests in fifth grade of child $i$ belonging to cohort $t$. $\alpha_{i}$ are city district fixed effects ${ }^{2}$ and $\lambda_{t}$ are cohort fixed effects. $D_{i, t}$, our variable of interest,

[^0]is a dummy variable equal to 1 if child $i$ lives in a city district with free ASP the year they start school. ${ }^{3} X_{i}^{\tau}$ is a vector of covariates measured at year $\tau$, i. e. when the child is age five, and indicates the socioeconomic characteristics of the individual (gender and birth quarter) and parental characteristics (whether mother and father is born abroad, mother's country of origin and parental education). $\varepsilon_{i, t}$ is the error term with conditional expectation zero. Standard errors are clustered at the city district level, accounting for dependency within city district. The main outcome variables are standardized results on national tests in 5 th grade in reading and mathematics.

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 $(\gamma)$ and post-treatment lags $(\delta)$ relative to the year before treatment, the omitted variable in the regression. Finding leads that are not significantly different from the omitted period lends support to our common trends assumption. If lags are significantly different from the omitted period, then this suggests effects of treatment and also shows how treatment effects develop over time. ${ }^{4}$

Due to the time between treatment (first grade) and testing (fifth grade), only two treated cohorts have data available for post-treatment outcomes; 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 use children starting school 2010 and onward. Treatment for the first two expansions implies that the after school program is free during the first four school years. We need to pay particular attention to the three pilot schools where the program was implemented before the roll-out. In the analysis we simply exclude pupils that take their national test at these schools (the results are robust to this exclusion).

In the period 2006-2016, children in many of the same districts that later were the first to introduce free ASP had an offer of free part-time child care. In 2016 this became national policy, but only for low-income families. This is not in itself a challenge for our empirical strategy, but

[^1]we need to take this into account when we decide on how many cohorts to include in the preperiod. We want to secure that there are no potential breaks in the trend assumption, and we do this by excluding cohorts that were not affected by this policy. For pre-years we therefore use children starting school 2010/11 and onward.

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 the first expansion group (early adopter) becoming 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 and by presenting results from various combinations of treatment and control group comparisons. ${ }^{5}$

## Data

## 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 directly 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 2010 and onwards.

## Outcome Variables

The main outcome variable is compulsory national tests in the subject reading and mathematics, taken in the beginning of 5 th grade, i.e. the cohort starting school in 2016 is tested in the fall

[^2]of 2020. The outcome consists of a continuous variable that measures overall performance as a scaled score, 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' 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.

We show separate estimates 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 for whether the child was exempted from the test.

## Control Variables and Sub-sample Stratification

The background characteristics of the children and their families are measured for the year before the child start school to ensure that they are not endogenous to the 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. In addition, we include a first generation immigrant variable picking up if the child is born abroad as well. We also construct dummies for mother's continent of origin. The control for family income is the average income of the mother and father, in addition to indicators for whether the mother or father receives welfare benefits. For educational attainment, we construct dummies on whether the mother/father has finished high school or college, respectively. We study sub-samples by parental education, immigrant background (mother and father is born abroad) and whether family income is below $40 \%$ of median family income.

Table 2shows 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 (0.32) and both are higher than the comparison group (0.13). The share with an immigrant background from Asia and Africa (mother's continent of birth) is 0.39 and 0.13 in the 2016 treatment group, 0.18 and 0.11 in the 2017 treatment group and only 0.07 and 0.04 in

Table 2: Summary statistics

|  | Treated 2016 | Treated 2017 | Comparison |
| :--- | :---: | :---: | :---: |
| GIRL | 0.49 | 0.49 | 0.49 |
| IMMIGRANT | 0.54 | 0.32 | 0.13 |
| CONTINENT |  |  |  |
| ASIA | 0,39 | 0,18 | 0.07 |
| AFRICA | 0,13 | 0,11 | 0.04 |
| AMERICA OCEANIA | 0,02 | 0,02 | 0.02 |
| EUROPE | 0,46 | 0,69 | 0.87 |
| PARENTS EDUCATION |  |  |  |
| M FINISHED HIGH SChool | 0,56 | 0,72 | 0.88 |
| F FINISHED HIGH SCHOOL | 0,56 | 0,70 | 0.86 |
| M UNIVERSITY | 0,32 | 0,55 | 0.74 |
| F UNIVERSITY | 0,30 | 0,47 | 0.67 |
| M EDU UNKNOWN | 0,10 | 0,07 | 0.03 |
| F EDU UNKNOWN | 0,09 | 0,08 | 0.04 |

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.

## Trends in Outcomes

We begin with a visual inspection of trends in our outcome variables separately for treatment and control groups. Figures 3 a and 3 b 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 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. The trends are much less stable prior to 2013.

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

Figure 3: Trend in National Test Scores

first graders (2017) relative to the comparison group.
A second observation is the comparison of the two treatment groups in 2016, 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 is that children with immigrant background would be the ones to benefit the most from the policy. This is both due to their low relative enrollment prior to introducing free ASP, and because attendance is thought to benefit them especially from exposure to the Norwegian language, in which case we may expect to see stronger effects for reading than for mathematics. In Figures 4 a and 4 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.

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 some 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 sign of an effect of free ASP on the first cohort of treated first graders (2016) while tndhere is a slight relative relative increase in the second year where both treatment groups (2017) have treated first graders. We also 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 more jumpy in the pre-treatment period. In the post-treatment period, there is a slight decrease in the share exempt in 2016 for the treatment group relative to the comparison group, followed by a slight increase in 2017.

## Results

We next turn to the results from running our model specified in equation (1), presented in Table 3. Outcome variables are scaled scores, normalized to have mean 0 and standard deviation 1 for the entire time period, proficiency indicator and whether students are exempted from the test, for both reading and mathematics. We restrict the estimation window to the cohorts starting first grades in the years 2013-2017, where we observed stable pre-trends across outcomes. In order to make our results as transparent as possible, particularly in light of the new developments
in the difference-in-differences literature, we present various combinations of treatment and control group comparisons. We first estimate equation (1) with the traditional staggered two way fixed effects design, showing both the aggregate estimate ( $D^{2016 / 2017}$ ) and the estimate for each post-treatment period ( $D^{2016}$ and $D^{2017}$ ). Second, since treatment occurred close in time, we run the estimation as if both treatment groups were treated in the first year (nonstandardized estimation), ensuring that we always compare to the non-treated city-districts. The first year of treatment is then a combination of treated and non treated cohorts from the treatment groups, which would reduce our treatment effect, 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, giving us the full treatment effect in the absence of cohort heterogeneity. Third, we separately compare the 2016 treatment group and the 2017 treatment groups to the comparison group, again ensuring that we are not using early adopters as a comparison group for the later adopters. For the former group, we present results for the two post-treatment periods separately. Finally, we focus on the first year of treatment for the 2016 treatment group, using the 2017 treatment group as a comparison 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.

The results reflect what we observed when looking at trend figures - for most outcomes there is no measurable effect of introducing free ASP on national tests regardless of specification. There are a few significant effects, but these are not consistent across specifications. Results controlling for background characteristics (Appendix Table A.1) and using the entire pre-trend period (Appendix Table A.2) confirm our results. In particular, we see no significant effects neither for the two way fixed effect approach nor for the second year of the non-standardized estimation where we might expect full treatment effects. The one exception is whether students are exempt from the test where the second treated cohort in the 2016 treatment group seems to have a negative estimate. This is a positive sign, as it is an indication that more students are able to take the test, i.e. do not require special needs education or have sufficient knowledge in Norwegian to complete the test. However, this finding is not consistent across treated cohorts.

As mentioned previously, average effects could conceal important heterogeneity, and we are particularly interested in whether children with immigrant background benefit from the

Table 3: Main results


Table 4: Main results - Immigrants

|  |  | Reading |  |  | Mathematics |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Two way fixed effects |  |  |  |  |  |  |
|  | $D^{2016 / 2017}$ | -0.009 | -0.026 | -0.007 | -0.017 | -0.013 | 0.005 |
|  |  | (0.031) | (0.016) | (0.012) | (0.041) | (0.022) | (0.011) |
|  | $D^{2016}$ | -0.047 | -0.043* | -0.011 | -0.047 | -0.027 | 0.003 |
|  |  | (0.034) | (0.018) | (0.013) | $(0.044)$ | (0.023) | (0.012) |
|  | $D^{2017}$ | -0.010 | -0.034+ | -0.009 | -0.066+ | -0.022 | 0.003 |
|  |  | (0.038) | (0.016) | $(0.015)$ | $(0.033)$ | (0.028) | (0.011) |
|  | N | 9331 | 9331 | 10208 | 9343 | 9343 | 10329 |
|  | Non-standardized estimation |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.098+ | -0.061* | -0.009 | -0.107* | -0.045 | -0.011 |
|  |  | (0.053) | $(0.024)$ | $(0.012)$ | (0.050) | $(0.032)$ | $(0.013)$ |
|  | $D^{2017}$ | $0.072+$ | 0.003 | -0.018 | -0.006 | -0.017 | -0.007 |
|  |  | $(0.035)$ | $(0.014)$ | $(0.012)$ | (0.039) | (0.031) | $(0.012)$ |
|  | N | 9331 | 9331 | 10208 | 9343 | 9343 | 10329 |
|  | 2016 treatment group and comparison group |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.118+ | -0.076* | -0.008 | -0.109+ | -0.042 | -0.004 |
|  |  |  |  | $(0.016)$ | (0.050) | (0.033) | $(0.014)$ |
|  | $D^{2017}$ | $0.073+$ | 0.001 | -0.017 | -0.013 | -0.010 | -0.008 |
|  |  |  |  | $(0.012)$ | (0.039) |  |  |
|  | N | 6861 | 6861 | 7428 | 6833 | 6833 | 7491 |
|  | 2017 treatment group and comparison group |  |  |  |  |  |  |
|  | $D^{2017}$ | 0.086+ | 0.017 | -0.017 | 0.034 | -0.015 | -0.000 |
|  |  |  |  |  |  |  |  |
|  | N | 4757 | 4757 | 5236 | 4782 | 4782 | 5325 |
|  | 2016 treatment group and 2017 treatment group |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.059* | -0.017 | -0.017 | -0.105* | -0.041 | -0.032* |
|  |  |  |  |  |  | (0.018) | (0.009) |
|  | N | 6173 | 6173 | 6797 | 6207 | 6207 | 6868 |
| Note: Standard errors in parentheses $+\mathrm{p}<0.10$, * $\mathrm{p}<0.05$ |  |  |  |  |  |  |  |

policy. Results are presented in Table 4 and show little signs of this being the case. Again, there are a few significant effects in both directions, but these are not consistent across treated cohorts or specifications. We also separate results by sub-samples using the two way fixed effects specification (aggregate and decomposed by post-treatment year) by immigrant background and family income (reported in Appendix Table A. 4 and A.5). None of the main estimates are significant and there does not seem to be a clear pattern of heterogeneity in treatment effects between subgroups.

Next, we present results for the event study specification in Figure 5 for all children and Figure 6 for children with immigrant backgrounds. The specification is based on the non-

Figure 5: Event study estimates of student outcomes, all children

standardized estimation with the cohort prior to treatment, 2015, as the excluded group. 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. Again we see that the first years in the pre-treatment period do not seem to follow the same trend for test scores, while the period from 2013 and onwards is more stable. 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. Figures A. 5 for all children and A. 6 for immigrant children in the appendix present results using the standardized approach where treatment occurs in year 0 for both groups and -1 is the excluded period, and findings are consistent.

Figure 6: Event study estimates of student outcomes, children with immigrant background




Above min. threshold math




## Conclusion

We have studied the roll-out of a free after school program in the municipality of Oslo. In the first wave, 34 schools were treated, followed by 20 new schools in the second wave. The take-up was substantial, raising enrollment rates rates from about 70 to $95 \%$ in many affected schools. Using unique 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 preprogram lower enrollment rate and as attendance possibly increased exposure to the Norwegian language. However, we find little support for enhanced academic performance for children with an immigrant background.

Former findings suggest that the quality of ASPs contribute to explaining positive effects (Kremer et al., 2015). The lack of effects on academic performance in our study raises the question of whether the quality of the program in Oslo is at a level that would lead to increased academic performance. Norwegian ASPs are more about free play than structured learning activities, and this may not be sufficient tools if the goal is to impact academic results. However, increased interaction with school peers could still matter for the social environment in school, and subsequently for student well-being.

We will continue to add future cohorts to the study, and importantly, to investigate outcomes related to questionnaires about student well-being and the more general learning and social environment at the school.

## References

Aizer, Anna, "Home alone: Supervision after school and child behavior," Journal of public economics, 2004, 88 (9-10), 1835-1848.

Baker, Andrew C, David F Larcker, and Charles CY Wang, "How much should we trust staggered difference-in-differences estimates?," Journal of Financial Economics, 2022, 144 (2), 370-395.

Bettinger, Eric, Torbjørn Hægeland, and Mari Rege, "Home with mom: the effects of stay-at-home parents on children's long-run educational outcomes," Journal of Labor Economics, 2014, 32 (3), 443-467.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event study designs: Robust and efficient estimation," arXiv preprint arXiv:2108.12419, 2021.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," Journal of Econometrics, 2021, 225 (2), 200-230.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," The Quarterly Journal of Economics, 2019, 134 (3), 1405-1454.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," American Economic Review, 2020, 110 (9), 2964-96.
_ and _ , "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey," Technical Report, National Bureau of Economic Research 2022.

Deshpande, Manasi and Yue Li, "Who is screened out? Application costs and the targeting of disability programs," American Economic Journal: Economic Policy, 2019, 11 (4), 213-48.

Felfe, Christina and Larissa Zierow, "After-school center-based care and children's development," The BE Journal of Economic Analysis \& Policy, 2014, 14 (4), 1299-1336.

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

Gormley, Todd A and David A Matsa, "Growing out of trouble? Corporate responses to liability risk," The Review of Financial Studies, 2011, 24 (8), 2781-2821.

Kremer, Kristen P, Brandy R Maynard, Joshua R Polanin, Michael G Vaughn, and Christine M Sarteschi, "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, 2015, 44 (3), 616-636.

Levine, Phillip B and David J Zimmerman, Targeting investments in children: Fighting poverty when resources are limited, University of Chicago Press, 2010.

Martínez, Claudia and Marcela Perticará, "Home alone versus after-school programs: the effects of adult supervision on child academic outcomes," International Journal of Educational Research, 2020, 104, 101601.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," arXiv preprint arXiv:2201.01194, 2022.

Schmitz, Laura, "Heterogeneous effects of after-school care on child development," Mimeo, March 2022. Working paper.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," Journal of Econometrics, 2021, 225 (2), 175-199.

## Appendix

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


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


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


Figure A.4: Trend in test excemptions, immigrant background
(a) Share excempted from reading test

(b) Share excempted from math test


Table A.1: Main results - With covariates


Table A.2: Results - Estimation period 2010-2017

|  |  | Reading |  |  | Mathematics |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Two way fixed effects |  |  |  |  |  |  | Exempted |
|  | $D^{2016 / 2017}$ |  |  |  |  |  |  |
|  |  | (0.024) | $(0.007)$ | $(0.005)$ | $(0.023)$ | $(0.008)$ | $(0.003)$ |
|  | $D^{2016}$ | -0.008 | -0.020* | -0.005 | -0.005 | -0.003 | -0.004 |
|  |  | (0.025) | (0.008) | (0.005) | (0.023) | (0.011) | (0.004) |
|  | $D^{2017}$ | 0.020 | -0.008 | -0.013* | -0.024 | -0.014 | -0.013* |
|  |  | (0.040) | (0.012) | (0.004) | (0.037) | (0.021) | (0.002) |
|  | N | 51168 | 51178 | 53835 | 51213 | 51223 | 54335 |
|  | Non-standardized estimation |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.043* | -0.021* | -0.006 | -0.069* | -0.034* | -0.007 |
|  |  | $(0.019)$ | $(0.007)$ | (0.005) | (0.026) | (0.015) | (0.005) |
|  | $D^{2017}$ | 0.036 | -0.003 | -0.012* | 0.014 | -0.008 | -0.013* |
|  |  | (0.040) | (0.012) | (0.004) | (0.030) | (0.015) | (0.002) |
|  | N | 51168 | 51178 | 53835 | 51213 | 51223 | 54335 |
|  | 2016 treatment group and comparison group |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.051* | -0.034* | -0.005 | -0.068* | -0.021 | -0.003 |
|  |  | $(0.021)$ | $(0.007)$ | (0.008) | (0.020) | (0.018) | (0.006) |
|  | $D^{2017}$ | 0.026 | -0.006 | -0.015* | -0.017 | -0.015 | -0.015* |
|  |  | (0.043) | (0.012) | (0.003) | (0.038) | (0.023) | (0.002) |
|  | N | 40304 | 40310 | 42145 | 40276 | 40282 | 42463 |
|  | 2017 treatment group and comparison group |  |  |  |  |  |  |
|  | $D^{2017}$ | 0.036 | -0.004 | -0.009 | 0.048 | 0.003 | -0.009+ |
|  |  | (0.045) |  |  |  |  |  |
|  | N | 37754 | 37764 | 39439 | 37778 | 37788 | 39812 |
|  | 2016 treatment group and 2017 treatment group |  |  |  |  |  |  |
|  | $D^{2016}$ | -0.085* | -0.033* | -0.009 | -0.062* | -0.029+ | -0.003 |
|  |  |  |  |  |  |  | (0.005) |
|  | N | 25398 | 25403 | 27180 | 25494 | 25499 | 27460 |
| Note: Standard errors in parentheses $+\mathrm{p}<0.10$, ${ }^{*} \mathrm{p}<0.05$ |  |  |  |  |  |  |  |

Table A.3: Main results - Immigrants with covariates


Table A.4: Immigrant background

|  | Reading |  |  |  | Mathematics |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Score | Proficiency | Exempted | Score | Proficiency | Exempted |  |
| Without immigrant background |  |  |  |  |  |  |  |
| $D^{2016 / 2017}$ | 0.024 | 0.007 | 0.004 | 0.021 | 0.015 | 0.000 |  |
|  | $(0.037)$ | $(0.010)$ | $(0.005)$ | $(0.027)$ | $(0.010)$ | $(0.004)$ |  |
| $D^{2016}$ | 0.017 | 0.001 | 0.006 | 0.029 | 0.019 | 0.003 |  |
|  | $(0.034)$ | $(0.009)$ | $(0.005)$ | $(0.027)$ | $(0.013)$ | $(0.003)$ |  |
| $D^{2017}$ | 0.050 | $0.031+$ | -0.006 | -0.010 | -0.002 | -0.011 |  |
|  | $(0.060)$ | $(0.015)$ | $(0.007)$ | $(0.051)$ | $(0.026)$ | $(0.008)$ |  |
| N | 23797 | 23797 | 24511 | 23792 | 23792 | 24760 |  |
| With immigrant background |  |  |  |  |  |  |  |
| $D^{2016 / 2017}$ | -0.009 | -0.026 | -0.007 | -0.017 | -0.013 | 0.005 |  |
|  | $(0.031)$ | $(0.016)$ | $(0.012)$ | $(0.041)$ | $(0.022)$ | $(0.011)$ |  |
| $D^{2016}$ | -0.020 | -0.029 | -0.006 | -0.018 | -0.015 | 0.006 |  |
|  | $(0.033)$ | $(0.018)$ | $(0.012)$ | $(0.045)$ | $(0.023)$ | $(0.011)$ |  |
| $D^{2017}$ | 0.048 | -0.009 | -0.010 | -0.011 | -0.001 | -0.001 |  |
|  | $(0.036)$ | $(0.016)$ | $(0.013)$ | $(0.036)$ | $(0.030)$ | $(0.013)$ |  |
| N | 9331 | 9331 | 10208 | 9343 | 9343 | 10329 |  |

Table A.5: Low income

|  | Reading |  |  | Mathematics |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Score | Proficiency | Exempted | Score | Proficiency | Exempted |
| Above median family income |  |  |  |  |  |  |
| $D^{2016 / 2017}$ | $\begin{gathered} 0.033 \\ (0.034) \end{gathered}$ | $\begin{gathered} 0.001 \\ (0.009) \end{gathered}$ | $\begin{gathered} 0.007 \\ (0.006) \end{gathered}$ | $\begin{gathered} \hline 0.033 \\ (0.025) \end{gathered}$ | $\begin{aligned} & 0.016+ \\ & (0.009) \end{aligned}$ | $\begin{gathered} 0.002 \\ (0.005) \end{gathered}$ |
| $D^{2016}$ | $\begin{gathered} 0.028 \\ (0.032) \end{gathered}$ | $\begin{aligned} & -0.000 \\ & (0.009) \end{aligned}$ | $\begin{aligned} & 0.011+ \\ & (0.006) \end{aligned}$ | $\begin{gathered} 0.038 \\ (0.022) \end{gathered}$ | $\begin{aligned} & 0.020^{*} \\ & (0.009) \end{aligned}$ | $\begin{gathered} 0.006 \\ (0.005) \end{gathered}$ |
| $D^{2017}$ | $\begin{gathered} 0.051 \\ (0.048) \end{gathered}$ | $\begin{gathered} 0.006 \\ (0.009) \end{gathered}$ | $\begin{aligned} & -0.006 \\ & (0.004) \end{aligned}$ | $\begin{gathered} 0.016 \\ (0.044) \end{gathered}$ | $\begin{gathered} 0.002 \\ (0.028) \end{gathered}$ | $\begin{gathered} -0.011^{*} \\ (0.003) \end{gathered}$ |
| N | 23324 | 23331 | 23944 | 23295 | 23302 | 23944 |
| Below median family income |  |  |  |  |  |  |
| $D^{2016 / 2017}$ | $\begin{aligned} & -0.026 \\ & (0.038) \end{aligned}$ | $\begin{aligned} & -0.024 \\ & (0.019) \end{aligned}$ | $\begin{aligned} & -0.017 \\ & (0.010) \end{aligned}$ | $\begin{aligned} & -0.020 \\ & (0.034) \end{aligned}$ | $\begin{aligned} & -0.000 \\ & (0.023) \end{aligned}$ | $\begin{aligned} & -0.003 \\ & (0.009) \end{aligned}$ |
| $D^{2016}$ | $\begin{aligned} & -0.039 \\ & (0.037) \end{aligned}$ | $\begin{aligned} & -0.031 \\ & (0.018) \end{aligned}$ | $\begin{aligned} & -0.017 \\ & (0.011) \end{aligned}$ | $\begin{aligned} & -0.024 \\ & (0.038) \end{aligned}$ | $\begin{aligned} & -0.004 \\ & (0.026) \end{aligned}$ | $\begin{aligned} & -0.002 \\ & (0.009) \end{aligned}$ |
| $D^{2017}$ | $\begin{gathered} 0.053 \\ (0.058) \end{gathered}$ | $\begin{gathered} 0.021 \\ (0.030) \end{gathered}$ | $\begin{gathered} -0.018+ \\ (0.009) \end{gathered}$ | $\begin{gathered} 0.005 \\ (0.037) \end{gathered}$ | $\begin{gathered} 0.021 \\ (0.022) \end{gathered}$ | $\begin{aligned} & -0.003 \\ & (0.010) \end{aligned}$ |
| N | 8193 | 8193 | 9073 | 8244 | 8244 | 9216 |

Figure A.5: Event study estimates of student outcomes, all children







Figure A.6: Event study estimates of student outcomes, children with immigrant background



Above min. threshold math





[^0]:    ${ }^{1}$ See Roth et al. (2022) and De Chaisemartin and d'Haultfoeuille (2022), for reviews covering the advances in the recent different-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)
    ${ }^{2} \mathrm{We}$ 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.

[^1]:    ${ }^{3}$ For the case with a single treatment time period, the model can be expressed as $Y_{i, t}=\alpha+\lambda_{t}+\beta D_{i}+\lambda\left(D_{i} \times\right.$ Post $\left._{t}\right)+\eta X_{i}^{\tau}+\varepsilon_{i, t}$ 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.
    ${ }^{4}$ 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).

[^2]:    ${ }^{5}$ This is comparable to creating "clean" $2 \times 2$ data sets prior to combining estimates in the stacked regression approach, see Cengiz et al. (2019); Deshpande and Li (2019); Gormley and Matsa (2011); Baker et al. (2022)

