Impact of funding targeted pre-school interventions on school readiness: Evidence from the Netherlands

Yusuf Emre Akgunduz

Suzanne Heijnen

Impact of funding targeted pre-school interventions on school readiness: Evidence from the Netherlands[*]

Yusuf Emre Akgunduz[†] Suzanne Heijnen[‡]

Abstract

We analyze the effectiveness of the early childhood programme (ECP) in the Netherlands. The programme is designed for 2.5 to 4 year olds from disadvantaged backgrounds. 37 municipalities received an additional subsidy to expand ECP programmes, which allows us to analyze the effects of the programme within a difference-in-difference-in-differences framework. Most children first enroll in primary schools at age 4 in the Netherlands, but pupils begin to learn reading and mathematics in grade 3 at age 6. We use grade repetition constructed from school registry data from 2008 to 2015 in the first two grades as an indicator of school readiness. Our results show significantly lower grade repetition rates for targeted boys who are in regions that receive the subsidy. Grade repetition drops by 1 to 3 percentage points from a mean of 10.5 percent for the disadvantaged group targeted by the programme.

JEL codes: C21, I28, I21, J13
Keywords: pre-school funding, early childhood, grade repetition, DDD

1 Introduction

The achievement gap between children from different socioeconomic groups forms before they even start primary school and persists later in life (Bradbury et al., 2015). Early childhood interventions have become an attractive policy tool to narrow the achievement gap at the start, especially since cognitive and non-cognitive skills appear to be most affected by investments in early years (Kautz et al., 2014). Evidence from small scale experiments support the use of early childhood interventions to close the achievement gap (Campbell et al., 2002; Heckman et al., 2010; Nores and Barnett, 2010). The success of small scale intervention studies has been followed by the development of larger scale targeted programmes such as Head Start in the United States and SureStart in England (Raver et al., 2008; Melhuish et al., 2008). The Netherlands has had a long standing policy of improving educational equality with a focus on 8th grade test scores and secondary education. In 2000, the early childhoood programme (*Voor- en Vroegschoolse Educatie*) was instituted, which is at least partly based on Head Start, with the aim of ameliorating disadvantages (particularly in language) at an early age. While most continental European and Nordic countries follow a rather universalist approach to early childhood education, the Netherlands explicitly targets disadvantaged children through the early childhood programme (henceforth ECP) programme.

This paper evaluates the impact of the Dutch ECP programme on school readiness. We attempt to find the causal effects, if any exist, of the ECP interventions on grade repetition of children in grades 1 and 2. The Dutch schooling system starts at an early age: 4. The ECP programme is aimed at children aged 2.5 to 4 prior to the start of primary school. The first two grades of primary school are equivalent to kindergarten (*kleutergroepen*) and grade 3 (when children are aged 6) is when the educational curriculum with reading and mathematics begins. Therefore, there is a large stepping stone between grades 2 and 3 and children who are not yet ready for schooling spend an additional year in the first two grades. The jump in intensity from grade 2 to 3 is also visible in grade repetition rates. The rate in the first two grades is 6% and 10% for disadvantaged children that the ECP programme targets.

To study the impact of ECP interventions, we exploit an ECP subsidy allocated to 37 Dutch municipalities in 2012 and 2013. The ECP budget in the Netherlands is divided across municipalities according to the number of disadvantaged children in primary school per municipality. In 2012 and 2013, 37 of the larger municipalities (M37) received additional funds equal to approximately 50 percent of their original funds to increase ECP quality and availability. The municipalities are then responsible for supplying ECP

through preschools and daycare centers. We exploit the targeted funding of municipalities in the Netherlands and the targeting of the programme for disadvantaged groups to estimate the effects of the ECP programme using a difference-in-difference-in-differences (DDD) strategy. By using grade repetition as the primary outcome variable, we are able to use administrative data provided by Statistics Netherlands and collected for all pupils in grades 1 and 2. The dataset is available for the years between 2008 and 2015 and each new cohort is comprised of approximately 180,000 pupils.

Our results show that grade repetition for targeted children was significantly reduced in municipalities that received the subsidy. Depending on the specification and sample, the effects range from 0.8 to 2.8 percentage points, which is equivalent to approximately 8 to 27 percent of average grade repetition rates for the targeted group. The effects are only for boys who have higher average repetition rates. The effects also appear to be somewhat weaker for pupils in the largest four cities. We further find that there was a significant increase in the availability of the ECP programme in the treated municipalities.

Previous research in the Netherlands has found negative or insignificant correlations between ECP attendance and performance in grade 2 test scores (Bruggers et al., 2014). Fukkink et al. (2015) provide a meta-analysis of previous Dutch studies of the effectiveness of ECP and find no significant effects on child development. However, finding the causal effect of early intervention programmes is challenging. Since children are selected into ECP due to potential development deficiencies, programme attendance may be correlated with negative outcomes, which would lead to biased estimates from correlational studies. Our analysis extends the literature to identifying the causal effects of ECP. Our broader contribution is to the literature on the impact of national targeted pre-school programmes, a literature which consists thus far almost entirely of analyses of Head Start. The literature on Head Start shows generally favourable effects, but the results are still mixed. The most recent national randomized study, the Head Start Impact Study (HSIS), showed positive effects on cognitive and non-cognitive outcomes at preschool age, but these effects appear to have faded out in the follow-up study in grade 2 (Puma et al., 2005, 2012). Deming (2009) shows that the effects fade out for test scores but return and are significantly positive again in an index of long-term outcomes including high school completion, college attendance, crime, early pregnancy and health status. In extremely disadvantaged areas, additional Head Start funding was also shown to have positive effects by Ludwig and Miller (2007) on long term education and health outcomes.

The remainder of the sections are organized as follows. Section 2 explains the most relevant policies and institutional background in the Netherlands. Section 3 discusses the methodology used to estimate the effects of ECP attendance and school subsidies. Section

4 introduces the datasets that we use. Section 5 provides the main results. Section 5 also presents the results from alternative specifications and placebo tests. Section 6 concludes.

2   Institutional background

The Dutch ECP programme consists of two parts: pre-school and early schooling. The pre-school programme is provided by daycare centers and playgroups starting at age 2.5 or 3. The early schooling portion of the programme is offered by primary schools at ages 4 and 5. The budget for the pre-school programme in 2012 was almost 7 times (€261 million) that of the early school programme (€50 million) and this paper focuses on the effects of the pre-school ECP programme.

The pre-school ECP programme was formalized in its current form in 2000. The design and contents of the programmes for children aged 2.5 to 4 are based on Head Start. The foremost goal is language development at early ages but development of calculation, motor and socio-emotional skills are also mentioned in the official programme documents. The programme consists of 10 to 12 hours of center care. The municipalities are responsible for supply. The national ECP budget is divided among the municipalities according to the school weighting system, which is also used to determine funding for schools. The school weights are calculated at the school level according to the number of disadvantaged pupils in the school.[1] To determine the share of the budget a municipality receives, a municipal weight score (the sum of the weighting scores of each school in that municipality) is calculated for each municipality. The national budget allocation has not been updated for changes in the municipal weight scores since 2009.

Municipalities differ not only in the budget they receive but also in the allocation and targeting. All municipalities have programs to encourage the use of ECP for disadvantaged groups by making it cheaper (or free of cost) and providing information, but the definition of disadvantaged groups differ. The common denominator is the same weighting system used to determine school subsidies and municipal weight scores. The targeted children are therefore those with parents who have at most early vocational training but some municipalities may further target children whose home language is not Dutch.[2] Administrative data on take-up is lacking, possibly because the programme is organized by municipalities rather than at the national level. The programmes are offered both at municipality run preschools and private daycare centers, which receive subsidies from the municipalities

---

[1]The exact formula is 1.2 times the number of highly disadvantaged pupils plus 0.3 times the number of moderately disadvantaged pupils minus the total number of pupils times 0.06.

[2]Although schools and municipalities both collect information on parents' education level for the ECP programme and the school weight scores, it is very likely that not all children who qualify are identified.

to offer the programme. By 2015, there were approximately 110,000 registered places in centers with ECP, which corresponds to around 35% of the cohort in grade 1 and 2. The actual number of ECP places is however likely to be lower since not all places in centers with ECP are designated for the ECP programme. The Ministry of Education reported that the number of ECP places was approximately 53,000 at the end of 2012. However, the ministry estimate excludes ECP places in 38 municipalities and the take-up rate is unknown.

Figure 1: Municipality locations
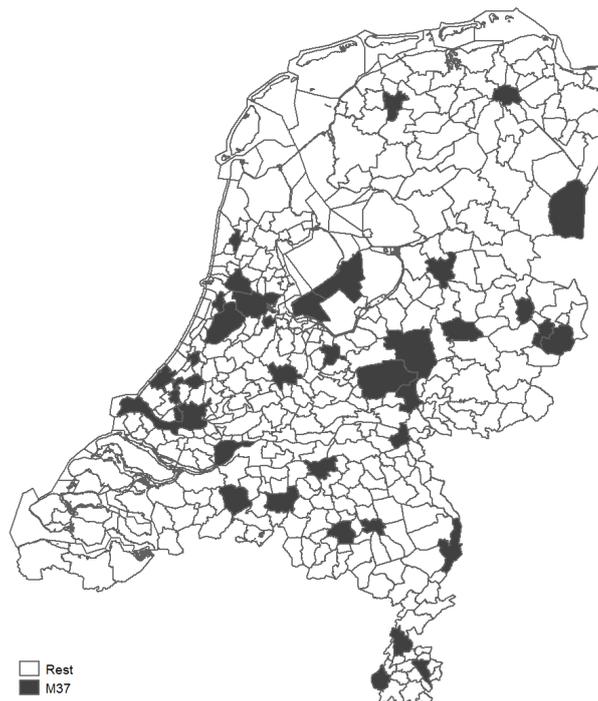
The budget of the ECP programme increased in the years following 2010 with the Law on Disadvantaged Education (*Wet OKE*). The budget was increased from 200 to 260 million euros in 2011. In 2012, and a further 70 million euro subsidy was made available for 37 large municipalities (M37) in the Netherlands. The M37 municipalities are the relatively central municipalities in all regions of the Netherlands and their locations are shown in figure 1. The additional amount for the 37 municipalities was raised further to 95 million euros in 2013. The subsidy is divided according to the municipal weight scores of municipalities. Per point, an additional subsidy of 1300 euros is paid on top of the 3000 euros per weighting score available through the main fund. As such, the treatment is equivalent to around 50% of the original budget, although there is some

Figure 2: Subsidy amounts per year (in millions)
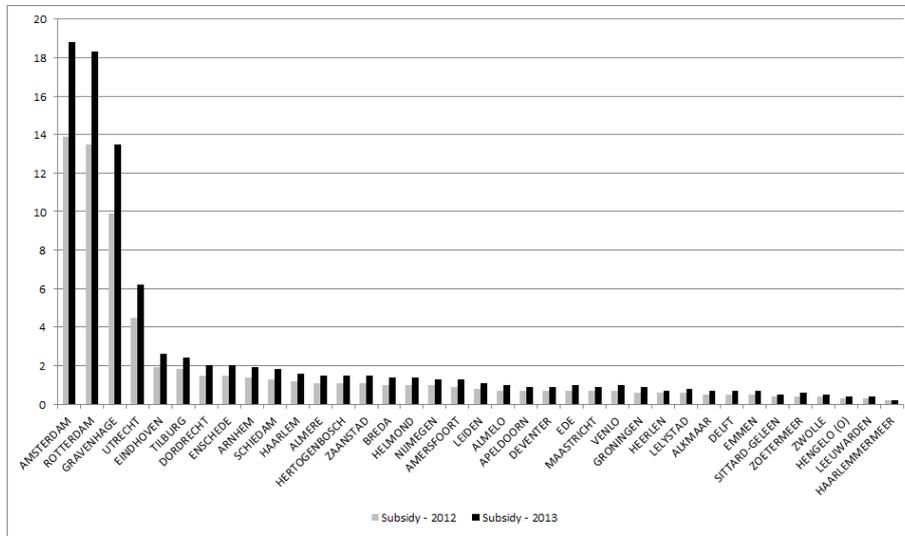
variation in the actual amounts. The absolute amounts for each municipality are presented in figure 2. Alongside the subsidy, a new monitoring programme with periodic reports was implemented to track the quality and the supply of the ECP programme in subsidized municipalities (OCW, 2014).

Figure 3: Subsidy amounts and population
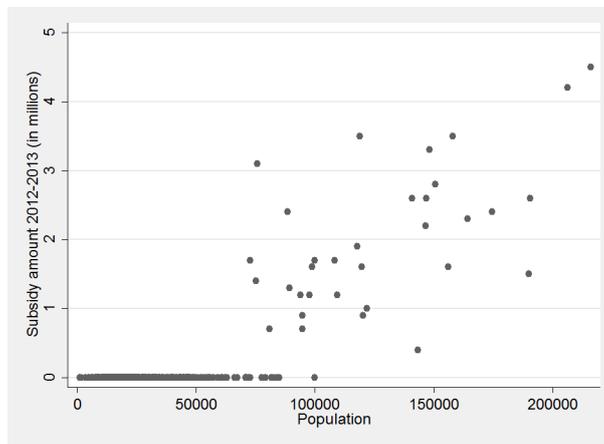
The subsidy was made available for M37 municipalities that are larger than the average in the Netherlands. Figures 3 and 4 plot the summed subsidy amount received by the municipality in 2012 and 2013 against the municipality's population and municipal weight

Figure 4: Subsidy amounts and municipal school weight
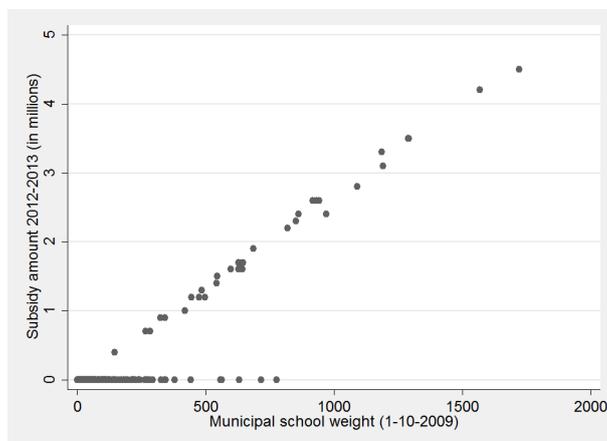
score.[3] While not visible in the figure, more than 50 out of 416 municipalities have a municipal weight score of 0 and receive no part of the ECP budget at all. We make use of this quasi-discontinuity in our analysis by limiting the sample to municipalities with a score of more than 0 and excluding municipalities with very large scores.

Figure 5: Grade repetition rate in Dutch primary schools 2012/2013
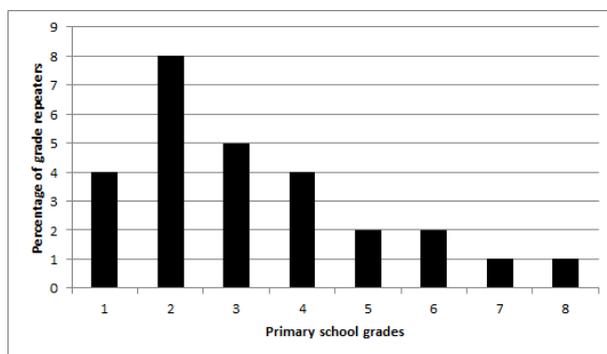
Our primary aim is to determine the effect of this additional subsidy on grade repetition in the first two years' of primary school. School starting age is 4 and a child can start school when turning 4 regardless of the month of the year with the exception of summer months. The first two grades of primary school are called toddler grades (*kleutergroepen*) and are comparable to kindergarten in other systems. In fact, until 1985, these first two grades were seperate from primary school and were simply named lower school or pre-school. Many schools do not distinguish between grades 1 and 2 and the groups are mixed. The

---

[3]We exclude the 4 largest cities from the figure to make the amounts more visible.

primary definition that we use for repetition in grades 1 and 2 is that a child spends three or more years in these grades instead of two. Children start structured reading and mathematics courses in grade 3. As a result, there is a large rate of grade repetition at grade 2 as can be seen in figure 5. Grade repetition in this context provides a general indicator for school readiness since normal schooling starts at grade 3.

Since the subsidy increase is first made available in 2012, the 2012 cohort of pupils would only be partially treated by the subsidy. Even those children who begin school in September can only benefit from the effects of the subsidy for 9 months while the total programme is potentially for up to 18 months, though limited participation is possible. We may therefore interpret the effects in 2012 only as partial treatment effects. However, the cohort that is first observed in primary school in 2013 should be largely exposed to the effects of the subsidy. If we take January 2012 as the starting point of the subsidy, it will have been in place for around 21 months prior to October 2013.

3 Methodology

The impact of ECP attendance can be estimated using OLS as in equation 1, where ECP_i is a dummy variable indicating whether a child has attended a ECP programme and X_i are a list of control variables for the child's school and family characteristics. However, the estimated coeffient α_1 may be biased due to unobserved factors that affect both ECP attendance and school performance. The children in ECP are by definition those from more disadvantaged backgrounds and there is likely to be selection of disadvantaged children who are doing particularly poorly. Children who attend ECP may then have unobserved disadvantages that would likely also affect performance in language and mathematics tests, leading to an underestimate of α_1.

$$y_i = \beta_0 + T_t + X_i'\beta + \alpha_1 ECP_i + u_i \tag{1}$$

To avoid the potential endogeneity problem, we use a difference-in-differences strategy to identify the effects of the extra ECP subsidy given to 37 Dutch municipalities in 2012 and 2013. Our emprical approach is to compare the outcomes for children from municipalities that receive the subsidy with children from other municipalities. The resulting estimate gives the effect of the additional funds on the targeted group rather than the effect of ECP attendance. The estimated regression equation is presented in equation 2. We regress whether a pupil i of cohort t in municipality j had grade repetition, y_{ijt} on year fixed effects, T_t, municipality fixed effects, R_j, individual characteristics X_{ijt} and a variable indicating the amount of the subsidy a region s will receive S_s at time p. Since the subsidy

8

amount is heavily dependent on the size of the municipality, we define S as the amount of subsidy per pupil in the municipality that is being phased in during that year. The parameter of interest is then γ_1 which will give the DD estimate of the effect of being in a region that receives the subsidy.

$$y_{ijt} = \beta_0 + T_t + R_j + X_i'\beta + \sum_{p=2012}^{2013} \gamma_{1p} S_{sp} + u_{ijt} \tag{2}$$

We can further exploit the targeting of the programme since not all pupils in subsidized municipalities are targeted by the programme. The ECP programme is explicitly targeted towards children with a disadvantage, the same disadvantage that is used to calculate the weights schools have. We can therefore use a difference-in-difference-in-differences (DDD) specification to estimate the effects of the subsidy on targeted regions for targeted disadvantaged pupils. To estimate the DDD model, we include interaction effects to the main model in equation 3 for pupils with a weight status alongside the pairwise interactions between the subsidy amounts and weight status and weight status and year fixed effects. Since there are two potential treatment years and the subsidy amounts differ between the years, we include pairwise and triple interactions of weight status with both 2012 and 2013 subsidy amounts. The DDD estimate of the effect of being a targeted pupil in a region that receives the subsidy is now given by the coefficient γ_4. We cluster the standard errors in all regressions at the municipality level to avoid overrejection of the null hypothesis (Bertrand et al., 2004).

$$y_{ijt} = \beta_0 + T_t + R_j + X_i'\beta + \gamma_0 W_{it} + \gamma_1 W_{it}T_t + \gamma_2 W_{it}S_s + \sum_{p=2012}^{2013} \gamma_{3p} S_{sp} + \sum_{p=2012}^{2013} \gamma_{4p} S_{sp} W_{it} + u_{ijt}$$
$$\tag{3}$$

Since the subsidy was only provided to large municipalities, that are central in their regions, there is a significant difference in the size and municipal weight scores of municipalities that receive the subsidy and those that do not. While there is no clear cut-off point seen in figure 3, we can use a quasi-regression discontinuity strategy by exploiting the targeting of the subsidy to only municipalities with high municipal weight scores (Lee and Lemieux, 2010). Our starting point is to set the sample to municipalities that receieve at least some of the ECP budget by having a minimum municipal weight score of 1. The sample is then limited further by excluding municipalities with very high and very low school weights. Selection based on weights also makes municipality sizes more similar between control and treatment groups. The municipalities of the 4 big cities (Amsterdam, Rotterdam, the Hague, Utrecht) have especially high municipal weight scores compared

to rest of the municipalities in the Netherlands. In section 5, we further test the validity of the estimates by estimating effects from placebo reforms that are assumed to happen in 2011 or in municipalities with high municipal weight scores that are not among the M37 municipalities.

4 Data

Our primary data source is the student registry database supplied by the Statistics Netherlands for the years between 2008 and 2015. The dataset is generated by compiling information provided by all primary schools in the Netherlands, who report the students enrolled in each grade along with their social security number, self-reported ECP use and weight score. The compilation is made on the 1st of October in each year. We merged this dataset with municipal residence data using the social security number of the pupils to get information on their date of birth, ethnicity and gender. A final dataset used is the Registry of Daycare centers and Playgroups (LRKP) in the Netherlands, which reports on whether a center offers ECP and the number of available places at the center, from which we calculate the number of potential ECP places on the 30th of January for each year. The summary statistics for the main variables used in the analysis are presented in table 1. The summary statistics for the restricted sample excluding small municipalities and the big 4 cities are shown in the Appendix table A1. In the main sample, treated municipalities have larger cohort sizes and higher proportions of pupils with weight status and migrant background. Simple t-tests show these differences to be statistically significant. In the restricted sample, the differences remain statistically significant, but become considerably smaller in size.

Table 1: Summary statistics of the full sample

	Pre-Subsidy		Post-subsidy	
	Rest	M37	Rest	M37
Repeated	0.0524	0.0577	0.0561	0.0596
Male	0.5091	0.5058	0.5099	0.5078
Migrant background	0.0727	0.2383	0.0703	0.2186
Weight status	0.0891	0.1517	0.0801	0.1253
Municipality cohort size	447.698	3275.561	445.0054	3465.679
School size	279.503	355.9799	275.0434	362.2887
N	310,936	207,533	202,440	141,190

Grade repetition per cohort is constructed using three years of data and is defined as

observing a pupil three times in grades 1 and 2 on October 1st of each year. If a pupil is first observed in grade 1 or grade 2 at year t, we check whether they are observed again at year $t + 2$ in grade 1 or grade 2. If so, the pupil is marked as having repeated, since the final year of observation should be $t + 1$. Since we need at least two years of data after the initial observation to determine repetition and the final year for which the dataset is available is 2015, the final cohort that we include are pupils that begin primary school in 2013-2014. Likewise, both grades 1 and 2 can be where a pupil is first observed and we therefore need a year at $t - 1$ to determine if an observation at time t is indeed the first observation. Therefore, the starting cohort for the study is 2009 rather than 2008, which is the starting year of our dataset. A small proportion of children (5%) are moved to grade 3 after only a single year in grades 1 or 2, but the majority move to grade 3 after two years in the first two grades. A small proportion (0.5%) take more than three years to complete the first two grades, but we drop these observations in the regression analysis since multiple years of repetition cannot be identified in 2012 and 2013.[4]

We retrieve background information on pupils using the population registry data provided by Statistics Netherlands. Two background characteristics are especially strong predictors of school repetition: gender and month of birth. Boys tend to have a higher rate of repetition than girls and the gender gap has been relatively stable over time as seen in figure 6. A second predictor is the month of birth as seen in figure 7. Since our observation of children is always in October, children born in the winter and spring tend to have lower rates of grade repetition. Children can enroll in grade 1 at any point during the academic year. Therefore a pupil who turns 4 in November and enrolls in school may be in grades 1 and 2 for 2 year and 11 months and would not be registered as a repeater since he would not be observed in October for a third time. On the other hand, children who are born in July, August and September are more likely to be observed a third time since they will first be observed within a few months of starting school.

The large difference in the repetition rates by the month of birth is largely because children are allowed to enroll when they turn 4 and might therefore be enrolled prior to the start of the academic year. While we observe a cohort on October the 1st of a given year, pupils of that cohort may have started school as early as the last quarter of the previous year. The enrollment timing has implications for both the probability of repetition due to time spent in school and the amount of treatment children have been exposed to. In table 2, we define both the treatment duration and repetition probability of the potential

---

[4]Their inclusion does not seem to change the results significantly. The absolute value of the coefficients for the subsidy variables become smaller by around 0.00015 in the main models when these observations are included. These observations are still included in figure 8.

Figure 6: Repetition rates by gender over time
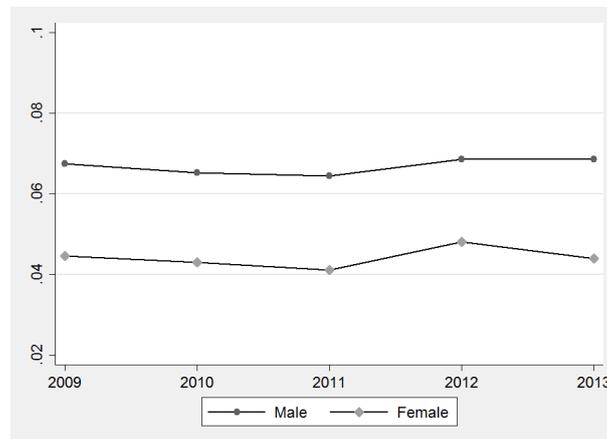
Figure 7: Repetition rates by month of birth
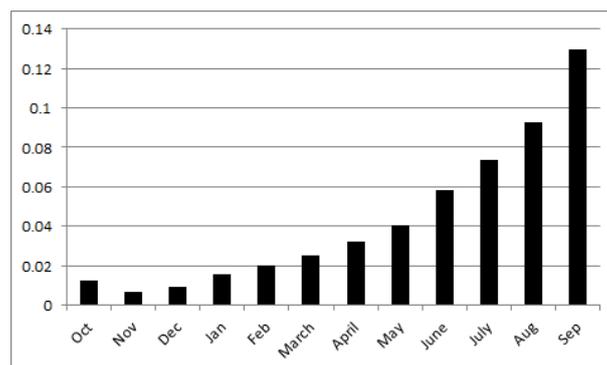
treatment cohorts of 2012 and 2013 by their quarter of birth. We seperate each cohort into four quarters of birth. The pupils in the first quarter are born between October to December of 2007 for the 2012 cohort and 2008 for the 2013 cohort. The pupils in fourth quarter are born in the months leading up to October: July to September. Column 3 shows the treatment duration for each group which is defined as the number of quarters after January 2012 that coincides with the period when the pupils were aged between 2.5 and 4. Clearly, we would expect the strongest effects to be seen for quarter 4 of the 2013 cohort since the repetition rate of quarter 4 is high due to their months of birth and they have been exposed to the longest treatment duration.

Table 2: Treatment and repetition probability of 2012 and 2013 cohorts

Observed	Quarter	Born	Treatment duration (quarters)	Repetition rate
2012 October	1	Oct-Dec 2007	0	Low
2012 October	2	Jan-Mar 2008	0	Low
2012 October	3	Apr-June 2008	1	Medium
2012 October	4	July-Sep 2008	2	High
2013 October	1	Oct-Dec 2008	3	Low
2013 October	2	Jan-Mar 2009	4	Low
2013 October	3	Apr-June 2009	5	Medium
2013 October	4	July-Sep 2009	6	High

Our identification has two sources. First is the targeting of the programme for pupils with weight status, whom we call targeted pupils, and the second is the regional targeting of the additional subsidy, which limits the treatment to targeted pupils in treated municipalities (M37). Figure 8 shows the average grade repetition rates across the years of interest for both targeted and non-targeted pupils in treated and control municipalities. The targeted and control regions appear to move together in the grade repetition rates of non-targeted pupils. However, there is more noise in the targeted population's repetition rates since the sample is relatively small, around 9-10% of all children. We check whether the identification is appropriate in section 5 by testing a placebo reform in 2011, where the gap between treated and control regions' repetition rates for the targeted pupils differ the most.

An important question is whether the ECP subsidies are likely to lead to a substitution between forms of center care. Kline and Walters (2015) and Elango et al. (2015) finds that the effects of Head Start were considerably larger for children who switched from home

Figure 8: Repetition rates in grades 1-2 over time in the full sample
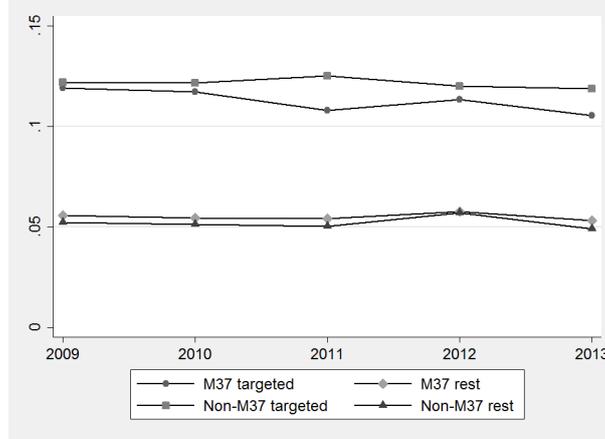
care compared to children who switched from a different preschool. The substitution effect extends to public pre-school expansions according to Cascio and Schanzenbach (2013) who find that expansion of public pre-school services led to substitution away from private pre-schools, especially for high income families. We checked the use of daycare centers for children with weight status one year before they were first observed in school by using the registry on childcare subsidies. Figure 9 shows that the pupils with weight status were much less likely to be in daycare centers than other children. Most children in playgroups with weight status are already likely to be in the ECP programme. The children who are induced to enroll in the ECP programme due to the subsidy will therefore likely be from daycare centers and home care. If the substitution is proportional to use of care type, at most 20% of children induced to enroll into ECP by the subsidy will be from daycare at least 80% will be from home care. The proportion switching from daycare centers is likely to be lower since some daycare centers already offered ECP services prior to 2012. We can then conclude that the substitution induced by the ECP subsidy will mostly be from home to center care in the Dutch context.

Figure 9: Daycare use among targeted and non-targeted children
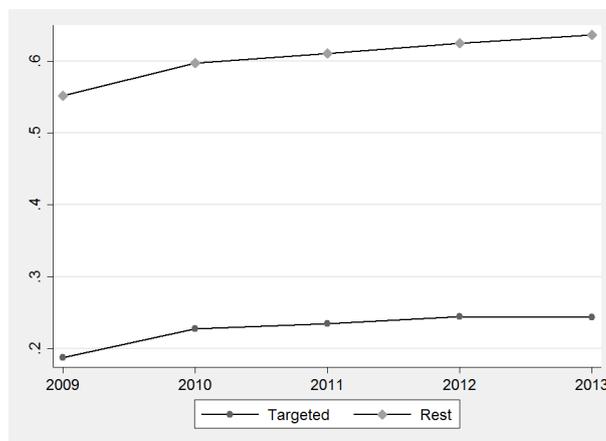
5 Results

5.1 Impact of the ECP subsidy on supply

Since the subsidy is earmarked for both quality and supply improvements, we first test whether there is an indication of greater supply of ECP in treated municipalities. Since ECP governance is highly decentralized, the sources on supply are not completely reliable. We use the Registry of Daycare centers and Playgroups (LRKP) dataset, which allows us to test whether the capacity in centers providing ECP is increased in M37 municipalities after the subsidy. Specifically, we look at the number of places in centers that offer the ECP programme on 30th of January for each year. Unfortunately, the LRKP registry only became compulsory for playgroups after 2012, which means we can only test the number of places in daycare centers. Furthermore, capacity in centers providing ECP does not imply that all places are reserved for ECP. In 2014, the Ministry of Education reported that the number of actual ECP places was approximately 55,000, around 32,000 of which were in treated municipalities. During the same year, the number of places in centers with ECP was approximately 100,000 in the LRKP dataset since all childcare places in a center offering ECP are counted as ECP places in the LRKP registry.

The results of the regressions based on the LRKP data are shown in table 3. Column 1 shows the results for the full sample, column 2 excludes municipalities with a weight score below 100, column 3 excludes the 4 big cities, while column 4 excludes both the municipalities below 100 municipal weight score and the 4 big cities. The effects appear to be consistently significant only in 2013, presumably because it takes some time for the subsidies to be used and distributed by municipalities to centers. Rather than using the

subsidy amount per child, we use a simple treatment variable that is 1 for M37 munici-palities and 0 for the rest in these estimates. If we assume the increase in places to be 250 per treated municipality, total increase can be said to be 9,250 within the 37 treated municipalities. We can take the ratio (0.55) between the number of places reported in the LRKP and the Ministry of Education to roughly convert the capacity of centers with ECP to actual number of ECP places. There can then be said to have been an increase of about 5,000 ECP places in daycare centers due to the subsidy. Since playgroups are not included in this calculation, the total number of new ECP places is likely to be even higher.

Table 3: DD estimates of the supply of ECP in daycare centers

	1	2	3	4
2012	109.3504*	91.6537	108.8508	91.1542
	(60.5612)	(62.3900)	(67.4221)	(69.2022)
2013	316.3261***	285.6232**	264.1746***	233.4717**
	(115.4988)	(118.0601)	(94.0653)	(96.8549)
N	1,436	424	1,420	408
Municipalities	359	106	355	102
Weight range	1-X	100-X	1-1721	100-1721

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. The treatment variable is an indicator variable of whether the municipality is an M37 municipality. All models include municipality and year fixed effects.

5.2   Impact of the ECP subsidy on grade repetition

Table 4 shows the results from the DDD specification in equation 3. Four models are fitted with varying treatment and control municipalities. The results show that the ECP subsidies have led to a significant reduction in the grade repetition of targeted children in treatment regions. Column 1 shows the results for the full sample. Column 2 excludes observations from municipalities that have a municipal weight score below 100. Column 3 excludes observations from the 4 big cities. Column 4 excludes both the municipalites below 100 municipal weight score and the largest 4 cities. The effect size varies between 0.6 and 2 percentage points per 1000 euro subsidies per child in the municipality in 2013. In line with the effects on ECP supply and the more limited time for treatment, there is no effect for the 2012 cohort. The effects seem considerably larger once the 4 big

16

cities are removed from the sample. The difference in effect size may be due to better implementation in small municipalities, the high availability of ECP in the 4 big cities prior to the subsidy or differences in grade repetition trends between the big 4 cities and the rest of the Netherlands. Since the average subsidy amount per child is approximately 1300 euros in the treated regions in 2013, the effect of the subsidies can be said to be between 0.8 and 2.8 percentage points. Given that the average repetition rate in 2011 for the targeted population was around 10.5%, the effect size is quite substantial.[5]

We also fitted the DD model in equation 2 and the results of the DD model are shown in the Appendix table A2. As might be expected, there is no overall effect on grade repetition if we simply compare pupils from M37 municipalities with pupils from the rest of the municipalities. Table A3 further shows the estimates from the interactions between the subsidy and weight status to have significantly negative effects within the DD. However, these estimates may be biased if children with a weight status have a different trend in grade repetition compared to other children.

Table 4: DDD estimates of the effects of 2012-2013 subsidies

	1	2	3	4
2012	0.0010	0.0018	0.0044	0.0070
	(0.0067)	(0.0077)	(0.0092)	(0.0108)
2013	-0.0059**	-0.0071*	-0.0164*	-0.0219**
	(0.0029)	(0.0036)	(0.0084)	(0.0096)
N	862,099	529,415	740,338	407,654
Municipalities	359	106	355	102
Weight range	1-X	100-X	0-1721	100-1721

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects.

Table 2 shows that the grade repetition is most likely in the last quarter of birth.

---

[5]While the national policy is to target children with weight status and the budget determined based on the number of children with weight status, most municipalities officially target more children than just those with weight status (Beekhoven et al., 2012). The most common additional criterium used for targeting is language spoken at home. When we exclude pupils with two non-Dutch parents and no weight status from the sample, the coefficient estimate in the full sample model is around 0.0005 larger.

Furthermore, children born between July and September 2009 are most exposed to the treatment. The repetition rate in 2011 for the targeted pupil born between July and September was approximately 21.7% as opposed to 10.5% in the full sample. We test whether the effects are indeed larger for pupils born in this quarter by restricting the sample to pupils born between July and September in all years. The results are shown in table 5. In line with our expectations, the estimates are 2 to 3 times as large as the effects found for the full sample. The same overall pattern holds as in the main results. 2012 effects are statistically insignificant while 2013 effects are statistically significant and larger once the 4 large cities are removed from the sample.

Table 5: DDD estimates for pupils born between July and September

	1	2	3	4
2012	0.0068	0.0087	-0.0047	-0.0016
	(0.0165)	(0.0188)	(0.0256)	(0.0294)
2013	-0.0130*	-0.0173*	-0.0395*	-0.0564**
	(0.0070)	(0.0089)	(0.0224)	(0.0257)
N	228,376	140,831	196,236	108,691
Municipalities	359	106	355	102
Weight range	1-X	100-X	0-1721	100-1721

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects.

Previous studies on early childhood interventions have shown significant differences by subgroups. There seems to be a clear difference between genders in the effects of national childcare reforms and Head Start. Studies of national childcare reforms in Spain and Norway find that the effects of greater availability of center care was particularly positive for disadvantaged groups and girls (Havnes and Mogstad, 2011; Felfe et al., 2012; Havnes and Mogstad, 2015). Meanwhile, Deming (2009) finds stronger effects for boys and Bitler et al. (2014) estimate larger effects for children with lower academic achievement from the targeted Head Start programme. Table 6 shows the results of DDD models fitted to the subsample of boys and girls in models 1 and 2 respectively. Column 3 shows the results for the subsample of children from schools that have an above average grade repetition

rate. We do not limit the sample based on the municipal school weight score for these models and use all municipalities with at least 1 municipal weight score. The results show that the effects are exclusively significant for boys and insignificant for girls. Compared to the effect size of 0.6 in table 4, the effects on the subsample of boys is almost twice as large. Similarly, the effects are twice as large as the main effects for schools where the grade repetition rate was above average, indicating that pupils in schools with below average performance are affected more positively from the ECP subsidy. The results of the targeted ECP programme for subsamples appears to be in line with the results from Head Start rather than national childcare reforms in Europe. The effects in targeted programmes seem to run through boys rather than girls and are stronger for schools with below average outcomes. A potential explanation is offered by Chetty et al. (2016), who show that the gender gap is reversed for children growing up in disadvantaged families. If Bertrand and Pan (2013) suggests boys are more likely to be affected by broken homes and family disadvantage, it seems plausible that they would be more likely to benefit from targeted early childhood interventions.

Table 6: Subsample DDD estimates

	1	2	3
2012	-0.0031	0.0050	-0.0011
	(0.0061)	(0.0089)	(0.0102)
2013	-0.0111***	-0.0007	-0.0111**
	(0.0043)	(0.0028)	(0.0051)
N	438,182	423,917	314,160
Municipalities	359	359	359
Weight range	1-X	1-X	1-X

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects. Model 1 uses provides the results for the subsample of males, model 2 for females and model 3 for pupils in schools with above average grade repetition rates.

Table 7: Placebo treatments in 2011

	1	2	3	4
2011	-0.0053	-0.0049	-0.0050	-0.0043
	(0.0045)	(0.0050)	(0.0089)	(0.0104)
N	518,469	316,170	446,833	244,534
Municipalities	359	106	355	102
Weight range	1-X	100-X	1-1721	100-1721

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects. We assume that the 2012 subsidy per child amounts were received in 2011.

5.3 Robustness tests

We report two sets of robustness tests to test the consistency of our results. We first perform a placebo test by assuming that the subsidies in 2012[6] were instead received by the municipalities in 2011. The results of the placebo test are shown in table 7. None of the specifications show significant effects from a hypothetical subsidy in 2011, which is reassuring for the main DDD models in table 3. On the other hand, the placebo reform for the simpler DD models show some significant effects as shown in the Appendix table A4. However, the significant effects in the DD models also disappear once municipalities with very low municipal weight scores and the largest four cities are excluded from the sample.

A secondary issue is whether the results are driven by the difference in the average municipal weight scores of the municipalities in control and treatment groups. We partially test this concern by limiting the sample and removing the 4 largest cities in table 3. However, the difference between the average municipal weight scores remain regardless and we cannot drop too many more treated municipalities with high municipal scores without making the treated sample too small.[7] However, we can test the validity of the

---

[6]2013 subsidy amounts differ very little from 2012 as seen in figure 2.

[7]If we drop 5 more of the remaining treated municipalities with highest municipal weight scores from model 4 in table 4, the DDD coefficient for 2013 (-0.0195) remains statistically significant at the 10% level. Dropping a further 6th municipality results in statistically insignificant effects but the point estimate remains negative at -0.0175 for 2013. At that point, the number of treated children seems too small to

Table 8: Placebo treatments in 37
untreated municipalities

	1	2
2012	-0.0120	-0.0156
	(0.0099)	(0.0134)
2013	0.0153	0.0103
	(0.0108)	(0.0143)
N	513,376	180,692
Municipalities	322	69
Weight range	1-X	100-X

*** p<0.01, ** p<0.05, * p<0.1. The
standard errors are clustered at the
municipality level. The treatment vari-
able is a dummy variable indicating
placebo subsidy receipt.

estimates in table 3 by assuming that the reform had happened in 37 municipalities with highest municipal weight scores that were not among the M37 municipalities which receive the ECP subsidy. We estimate DDD models where the placebo reform takes place in these 37 municipalities and the treatment variable is a dummy variable indicating the placebo subsidy receipt per municipality. The results are shown in table 8 and no statistically significant effects are found from the placebo reform. In fact, the point estimates in 2013 are positive rather than negative. The results seem to suggest that the effects are not driven by the differences in municipal weight scores between treatment and control regions.

A number of small variations can be made in the specification that we use to estimate the DDD models. We show results from four such specifications in table 7 using the full sample of municipalities with a municipal weight score above 0. Column 1 shows the results when municipality specific time trends are included. Column 2 replaces the subsidy per child variable with a dummy variable indicating subsidy receipt per municipality. Column 3 includes school fixed effects rather than municipality fixed effects as in the main model. Finally, column 4 shows the results of limiting the control group to municipalities that neighbor treatment municipalities. The coefficients from the 2013 subsidy are negative in all models and the sizes are similar to that in the main specification in table 3. The biggest difference is that the model with neighboring municipalities leads to larger

find statistically significant effects.

standard errors and smaller effects. The results might be pointing to spillover effects for smaller municipalities that are immediate neighbors of larger municipalities. However, the evidence for spillover effects remains tenuous since neighboring municipalities tend to be (often richer) suburbs of larger municipalities with significantly different populations.

Table 9: Alternative specifications and robustness tests

	1	2	3	4
2012	0.0012	0.0015	0.0016	0.0010
	(0.0068)	(0.0073)	(0.0067)	(0.0076)
2013	-0.0059**	-0.0105*	-0.0061**	-0.0042
	(0.0028)	(0.0063)	(0.0029)	(0.0034)
N	862,099	862,099	862,099	599,792
Municipalities	359	359	359	179

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects. Model 1 includes municipality specific time trends. Model 2 uses a dummy for treated areas rather than subsidy amounts per pupil. Model 3 adds school fixed effects. Model 4 limits the control group to municipalities neighboring treatment municipalities.

6 Conclusions

This paper analyzed the effect of a subsidy for an early childhood education programme targeted towards disadvantaged children. The literature on national childhood interventions is largely limited to Head Start despite the potentially large gains and tendency towards targeted programmes across OECD countries (Bennett et al., 2012; Elango et al., 2015). The Dutch case suggests that the results are positive also in Europe. The Dutch schooling system allows for a grade repetition outcome in the first two years of school that we use to measure school readiness since traditional schooling begins in grade 3. We find large effects on grade repetition and the results are robust to a variety of tests and specifications. Additional funds to the municipalities in the decentralized governance system of the ECP programme in the Netherlands appear effective at raising pupils' performance

at school.

Our results are contrary to the findings of the previous literature in the Netherlands, which finds insignificant effects from ECP attendance on child development (Fukkink et al., 2015). We expect that the difference is due to the selection issues in comparing children who attend ECP with those who do not. If more disadvantaged children are systematically selected into ECP programmes, positive effects may not be found even if some of the observables are controlled for.

Grade repetition is only one of the potential outcomes that ECP programmes may have an effect on. However, it is a costly policy tool in itself and discussion on the effectiveness and efficiency of grade repetition practices continue (Jacob and Lefgren, 2009; Manacorda, 2012). If we consider grade repetition and early childhood interventions to be substitutable catch-up mechanisms for disadvantaged pupils, early childhood interventions appear to save both time and money. Van Vuuren and Van der Wiel (2015) estimate that grade repetition costs around 500 million euros per year in primary and secondary education, which is equivalent to nearly 3% of the education budget in the Netherlands. If the decline is grade repetition is carried to future years and there are more positive effects on test scores and emotional development, we can expect the ECP policy to pay for itself.

References

Beekhoven, S., Jepma, I., Leseman, K.P., 2012. Landelijke Monitor Voor- en Vroegschoolse Educatie 2011. SARDES.

Bennett, J., Gordon, J., Edelmann, J., 2012. Early Childhood Education and Care (ECEC) for Children from Disadvantaged Backgrounds: findings from a European literature review and two case studies. Final Report for the Director–General. Brussels, Belgium: European Commission, Directorate–General for Education and Culture.

Bertrand, M., Duflo, E., Mullainathan, S., et al., 2004. How Much Should We Trust Difference-in-Differences Estimates? The Quarterly Journal of Economics 119, 249–275.

Bertrand, M., Pan, J., 2013. The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. American Economic Journal: Applied Economics 5, 32–64.

Bitler, M.P., Hoynes, H.W., Domina, T., 2014. Experimental Evidence on Distributional Effects of Head Start. Working Paper 20434. National Bureau of Economic Research.

Bradbury, B., Corak, M., Waldfogel, J., Washbrook, E., 2015. Too many children left behind: The US achievement gap in comparative perspective. Russell Sage Foundation.

Bruggers, I., Driessen, G., Gesthuizen, M., 2014. Vooren vroegschoolse voorzieningen, effectief of niet?/Early childhood facilities: effective or ineffective? Mens en Maatschappij 89, 117.

Campbell, F.A., Ramey, C.T., Pungello, E., Sparling, J., Miller-Johnson, S., 2002. Early childhood education: Young adult outcomes from the Abecedarian Project. Applied Developmental Science 6, 42–57.

Cascio, E.U., Schanzenbach, D.W., 2013. The Impacts of Expanding Access to High-Quality Preschool Education. Working Paper 19735. National Bureau of Economic Research.

Chetty, R., Hendren, N., Lin, F., Majerovitz, J., Scuderi, B., 2016. Childhood Environment and Gender Gaps in Adulthood. Working Paper 21936. National Bureau of Economic Research.

Deming, D., 2009. Early childhood intervention and life-cycle skill development: Evidence from Head Start. American Economic Journal: Applied Economics , 111–134.

Elango, S., Garca, J.L., Heckman, J.J., Hojman, A., 2015. Early Childhood Education. Working Paper 21766. National Bureau of Economic Research.

Felfe, C., Nollenberger, N., Rodríguez-Planas, N., 2012. Cant buy mommys love? Universal childcare and childrens long-term cognitive development. Journal of population economics 28, 393–422.

Fukkink, R., Jilink, L., Oostdam, R., 2015. Met een blink op de toekomst: Een meta-analyse van de effecten van vve op de ontwikkeling van kinderen in Nederland. Hogeschool van Amsterdam.

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

Havnes, T., Mogstad, M., 2015. Is universal child care leveling the playing field? Journal of Public Economics 127, 100–114.

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A., 2010. The rate of return to the HighScope Perry Preschool Program. Journal of Public Economics 94, 114 – 128.

Jacob, B.A., Lefgren, L., 2009. The effect of grade retention on high school completion. American Economic Journal: Applied Economics 3, 33–58.

Kautz, T., Heckman, J.J., Diris, R., Ter Weel, B., Borghans, L., 2014. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. Technical Report. National Bureau of Economic Research.

Kline, P., Walters, C., 2015. Evaluating Public Programs with Close Substitutes: The Case of Head Start. Working Paper 21658. National Bureau of Economic Research.

Lee, D.S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. Journal of Economic Literature , 281–355.

Ludwig, J., Miller, D.L., 2007. Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. The Quarterly Journal of Economics 122, 159–208.

Manacorda, M., 2012. The cost of grade retention. Review of Economics and Statistics 94, 596–606.

Melhuish, E., Belsky, J., Leyland, A.H., Barnes, J., of Sure Start Research Team, N.E., et al., 2008. Effects of fully-established Sure Start Local Programmes on 3-year-old

children and their families living in England: a quasi-experimental observational study. The Lancet 372, 1641–1647.

Nores, M., Barnett, W.S., 2010. Benefits of early childhood interventions across the world:(Under) Investing in the very young. Economics of education review 29, 271–282.

OCW, 2014. Monitor Kwaliteit van Voor- en Vroegschoolse Educatie in de 37 grote steden in 2013 en 2014. Technical Report. Ministerie van Onderwijs, Kultuur en Wetenschap.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., Downer, J., 2012. Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012-45. Administration for Children & Families .

Puma, M., Bell, S., Cook, R., Heid, C., Lopez, M., 2005. Head Start Impact Study: First Year Findings. Administration for Children & Families .

Raver, C.C., Jones, S.M., Li-Grining, C.P., Metzger, M., Champion, K.M., Sardin, L., 2008. Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. Early Childhood Research Quarterly 23, 10–26.

Van Vuuren, D., Van der Wiel, K., 2015. Zittenblijven in het primair en voortgezet onderwijs: Een inventarisatie van de voor- en nadelen. Centraal Planbureau.

Appendix

Table A1: Summary statistics of the restricted sample sample

	Pre-Subsidy		Post-subsidy	
	M37	Rest	M37	Rest
Repeated	0.0567	0.0598	0.0613	0.0636
Male	0.5051	0.5010	0.5071	0.5058
Migrant background	0.1616	0.1179	0.1490	0.1112
Weight status	0.1214	0.1190	0.1018	0.1061
Municipality cohort size	1567.0110	616.9307	1589.0550	618.0586
School size	352.9192	308.7196	353.3451	301.8235
N	108,637	135,897	72,055	91,065

Table A2: DD estimates of the effect of 2012-2013 subsidies on all pupils

	1	2	3	4
2012	-0.0027	-0.0023	-0.0060	-0.0058
	(0.0023)	(0.0031)	(0.0053)	(0.0068)
2013	0.0004	-0.0012	0.0023	-0.0004
	(0.0025)	(0.0032)	(0.0038)	(0.0046)
N	862,099	529,415	740,338	407,654
Municipalities	359	106	355	102
Weight range	1-X	100-X	1-1721	100-1721

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The standard errors are clustered at the municipality level. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects.

Table A3: Interaction effects in the DD model

	1	2	3	4
Main effects				
2012	-0.0027	-0.0023	-0.0060	-0.0058
	(0.0023)	(0.0031)	(0.0053)	(0.0068)
2013	0.0004	-0.0012	0.0023	-0.0004
	(0.0025)	(0.0032)	(0.0038)	(0.0046)
Interactions				
2012	-0.0123***	-0.0124***	-0.0072	-0.0101
	(0.0042)	(0.0046)	(0.0083)	(0.0085)
2013	-0.0112***	-0.0111***	-0.0136**	-0.0155**
	(0.0018)	(0.0020)	(0.0068)	(0.0070)
N	862,099	529,415	740,338	407,654
Municipalities	359	106	355	102
Weight range	1-X	100-X	1-1721	100-1721

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects.

Table A4: Placebo treatments in the DD model

	1	2	3	4
2011	-0.0053***	-0.0046**	-0.0046	-0.0040
	(0.0019)	(0.0022)	(0.0046)	(0.0056)
N	518,469	316,170	446,833	244,534
Municipalities	359	106	355	102
Weight range	1-X	100-X	1-1721	100-1721

*** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects, gender, migrant background and month of birth fixed effects.

Table A5: Extended DDD estimates of the effects of 2012-2013 subsidies

	1	2	3	4
Male	0.0228***	0.0229***	0.0232***	0.0236***
	(0.0007)	(0.0009)	(0.0006)	(0.0009)
Migrant background	0.0269***	0.0258***	0.0293***	0.0286***
	(0.0021)	(0.0024)	(0.0022)	(0.0027)
Weighted status	0.0637***	0.0682***	0.0601***	0.0628***
	(0.0032)	(0.0045)	(0.0033)	(0.0048)
2012 subsidy amount*weight status	-0.0384	-0.0361	-0.0051	-0.0082
	(0.0776)	(0.0752)	(0.0861)	(0.0867)
2013 subsidy amount*weight status	0.0189	0.0148	0.0008	-0.0003
	(0.0565)	(0.0545)	(0.0626)	(0.0630)
Subsidy*2012	-0.0027	-0.0026	-0.0062	-0.0066
	(0.0028)	(0.0037)	(0.0052)	(0.0067)
Subsidy*2013	0.0009	-0.0006	0.0039	0.0018
	(0.0025)	(0.0032)	(0.0034)	(0.0042)
2012*weight status	-0.0065	-0.0072	-0.0060	-0.0071
	(0.0044)	(0.0062)	(0.0047)	(0.0069)
2013*weight status	0.0024	0.0053	0.0057	0.0125*
	(0.0047)	(0.0066)	(0.0050)	(0.0073)
Subsidy*2012*weight status	0.0010	0.0018	0.0044	0.0070
	(0.0067)	(0.0077)	(0.0092)	(0.0108)
Subsidy*2013*weight status	-0.0059**	-0.0071*	-0.0164*	-0.0219**
	(0.0029)	(0.0036)	(0.0084)	(0.0096)
N	862,099	529,415	740,338	407,654
Municipalities	359	106	355	102
Weight range	1-X	100-X	0-1721	100-1721

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The standard errors are clustered at the municipality level. The treatment variables are subsidy per child in the municipality in thousands of euros. Additional controls are school size, municipality cohort size, school denomination fixed effects and month of birth fixed effects.