Assessing the effectiveness of early childhood programs in the Netherlands

Abstract:

In this paper I investigate the effectiveness of early childhood programs (ECP) in the Netherlands. By lack of a good experimental setting and instrumental variables, an alternative method, developed by Altonji, Elder & Taber (2005), is used. Because negative selection is likely to occur in ECP enrollment, I assume equal selection between observables and unobservables. ECP effects are assessed on cognitive skills and pupil characteristics in grade 2 and 5 in three steps. The main conclusion is that ECP-children have lower school results than non-ECP children. However, a large selection is shown in the ECP enrollment. It is important that previous studies are evaluated well, as most of them did not incorporate the selection effect well, but concluded that ECP is ineffective. Moreover, nothing can be said about the marginal development of the children. Strong conclusions about ECP effectiveness can therefore not be made.

August 23, 2016

Michelle Lim, studentnr. 415113tl

Erasmus University Rotterdam

Erasmus School of Economics

Department of Economics

Master's Thesis Policy Economics

Supervised by prof. Dinand Webbink

Co-reader: MSc. Matthijs Oosterveen

Contents

1. Introduction

A large body of literature shows that the early development of children has large and long-lasting effects (e.g. Heckman et al, 2010, Melhuish et al, 2015). Early childhood programs appear to have a major positive impact on school performance, earnings, future employment and crime rates. The evidence is restricted to early childhood programs in the US and of older cohorts. It is unclear whether these results can be extrapolated for European countries and to more recent cohorts, specifically for the Netherlands.

The Netherlands has special early childhood programs (ECP), which attempt to stimulate the child's language development in a playful way. This can be done by language games or by reading out. Other skills are stimulated as well, such as social-emotional skills and basic mathematics skills (CPB, 2016). An amount of approximately €254 million is devoted to the prevention of development delay for young children between the age of 2 to 5. This is part of the act 'Development Opportunities by Quality and Education' *(wet 'Ontwikkelingskansen door Kwaliteit en Educatie')* which was implemented on August 1, 2010. By means of this act, municipalities are obliged to offer preschool education to young children with a higher probability of developing a language disadvantage. Preschools and childcare centers can offer ECP if they meet special criteria. These criteria are set to guarantee the quality of the program (Cebeon & Regioplan, 2015).

Recent Dutch studies show mixed results. Two opposite conclusions about early childhood education recently reached the media: "*early language lessons help stimulating development*" (NRC Handelsblad, 2016) and "*the early childhood program does not add anything to the existing preschools and childcare centers*" (Brandpunt, 2015). The question that remains unanswered is: '*What is the effectiveness of the Early Childhood Program in the Netherlands?*'

This study investigates the effectiveness of ECP on cognitive skills and pupil characteristics, such as language, maths and reading skills, but also the effect on underachievement. Investigating the effectiveness of ECP is difficult because children are not randomly assigned to ECP. Therefore an alternative method, developed by Altonji, Elder & Taber (2005), is used. They assume equal selection on observable and unobservable variables. With a selection effect, there is a possibility that unobserved variables play a role in the way pupils perform at school, which can bias results. ECP is only available for certain types of children, and hence a selection effect is present. Not taking these selection effects into

account can bias results. For example, some characteristics of parents are important for the development of the child and are also likely to influence the skill level. However, these characteristics also have an influence on the decision whether the child is going to attend ECP or not. Without ECP, it is likely that a child whose parents would decide to enroll into ECP will perform better than the child who is not enrolled by its parents. If the first child follows ECP, the difference between the two children will become larger, as the first child gets extra attention in his/her development. Considering this, the reason that ECP-children are performing better is not because of following ECP, but due to differences in characteristics. ECP only attracts children whose parents value their child's development important. For this reason, I assess the effectiveness of ECP on different skills in second and fifth grade of primary school, while assuming equal selection of observable and unobservable variables. Section 4 provides a detailed explanation of the methodology.

I find that the majority of the ECP effects are negative. However, there seems to be a large selection effect. Taking the school results of all children into account, the children that followed ECP score significantly lower than the other children. When selecting on socio-economic status (SES), the differences are smaller and not always significant. The last step, assuming equal selection, shows even smaller differences. It is not guaranteed that the selection effect is fully filtered out by this study. Therefore it is still possible that the differences in school results are even smaller in reality. The main contribution of this study to the existing literature is the focus on the Netherlands and the use of more recent cohorts. Most importantly, it attempts to filter out the selection effect by dropping the conditional independence assumption and assuming equal selection of observables and unobservables. As far as I know, this is the first application to assessing the effectiveness of ECP.

The remainder of the paper is organized as follows. Section 2 introduces the early childhood education and care system in the Netherlands, and will primarily focus on the institutional part of the Dutch school system and ECP. Section 3 discusses the literature about ECP in and outside the Netherlands. In Section 4 the methodology will be explained. Section 5 gives detailed information about the data and a descriptive analysis. Section 6 will give the estimation results, both OLS estimations and the estimations that assume equal selection.

Section 7 will provide a discussion about the main results found in this study and some recommendations for future research. This paper concludes in Section 8.

2. Early Childhood Program: institutional

In the Netherlands different forms of early childhood education and care (ECEC) are available. They are either focused on daycare, education or a mix of both. Forms of daycare are available in the formal and the informal sector. In this paper, I will only pay attention to the formal sector. Here working parents send their offspring (age between 0-4) to a daycare center, with the main focus being taking care of the children while the parents are working. Preschools (*peuterspeelzalen*) offer children in the age 2.5 to 4 the possibility to enroll for 8 to 12 hours a week (generally 2 or 3 mornings). The focus of preschool is on social-emotional and cognitive development (Driessen & Doesburgh, 2003). The participation rate of ECEC is 83% in the Netherlands, which is high compared to other OECD-countries (*Organisation for Economic Co-operation and Development)* (OECD, 2015). The total amount of hours of participation for these children is low. On average children spend 2 days a week, about 16 hours, in a daycare center. Children in preschool only spend about 8 to 12 hours a week at school. Despite the high participation rate, a selection effect occurs in ECEC enrollment. Looking at socio-economic background, large differences are visible. For example, in the lowest income group 40 percent of the children are not involved in a form of ECEC. In the highest income group, this accounts for only 8 percent of the children (CBS, 2015).

In the Netherlands, children start learning the basic skills of reading and mathematics in grade 3. If children start with a language development delay, they are not ready to learn in third grade. This will cause an extra development delay in other cognitive skills. It is proven that pupils with a low language skill level start with a disadvantage. This does not disappear or diminish over the years (Driessen & Doesburgh, 2003). Driessen (2004) studied the disadvantage of migrant children and concluded that these children start with a language development delay of approximately two years in grade 3. This gap between migrant and non-migrant children is stable during primary school. Hence, it is better to prevent a disadvantage than to fight the disadvantage in later years.

Therefore, the Dutch government implemented special early childhood programs in 1998. ECP is a special and more intense program of 16 to 20 hours a week, with a stronger focus on the child's language development, meant for children with a high probability on development disadvantage. It is offered by daycare centers, preschools and kindergarten (grade 1 and 2 in primary school). By letting children enroll into ECP at a young age (starting at the age of 2), they will improve their cognitive skills and will start grade 3 without a

development delay (CPB, 2016). The main goal of ECP is to prevent disadvantages in primarily language skills. Besides the common reason that children have to start at a similar level in grade 3 to get the same opportunities, prevention is also very important, because of the early tracking system in the Netherlands. At the age of 12, pupils are tracked into a certain level of secondary education. This level already prepares for vocational or higher vocational education or university. This form of early tracking is heavily debated in the Netherlands. Especially if children do not get the opportunities to perform at their full potential because of a development delay in their early years, these children can become disadvantaged throughout their entire life. Moving up in the tracking system is possible, but it takes a lot of courage, time and money.

In the period 2007-2012, the Dutch Inspectorate of Education quantified the quality of ECP. Since then, a lot of improvement is seen in ECP (Inspectie van het Onderwijs, 2014). In the beginning, daycare centers and preschools had to see the advantages of ECP, such that they would implement ECP. After the programs were implemented, ECP had to be more familiar among parents to apply for these programs. Since 2010 the government put a stronger emphasis on the quality of ECP. This is formalized in the OKE-act (*Ontwikkelingskansen door Kwaliteit en Educatie)*. The aim of this act is to stimulate language development for young children and to improve and strengthen the quality of ECP. Hence, when the quality of teaching is lacking behind, the effect of participating in ECP will not be visible (Ministerie van OCW, 2013).

The responsibility of ECP is with the government and the municipalities. The government provide municipalities with financial resources, but the implementation of ECP is decentralized to the municipalities. They are responsible for providing high-quality ECP for all children with high risk on a language disadvantage. The total amount of money a municipality receives is based on the number of low-SES children. This is done by a weight scheme, looking at the educational background of the parents. Children with lower educated parents get a higher weight, which translates into more money for a municipality. This does not mean that these children can directly enroll in ECP. A municipality can target children in their own municipality, and can decide which children can participate in ECP. The main focus is that municipalities reach the children that are likely to start grade 3 with less developed language skills. Examples for these groups are children with non-Western parents or children with a low socio-economic status (low-SES). In most municipalities the children's health clinics have to

signal these children and inform the parents about the existence of ECP and encourage them to enroll. Schools and daycare centers that provide ECP have to comply to certain conditions: children have to enroll at an early age, between 2 and 2,5 years. Participation is 4 mornings in daycare centers and preschools and 5 days in kindergarten (grade 1-2 in primary school). The ratio teacher to pupil ratio is 1 to 8 and in grade 1 and 2 two teachers are necessary for 16 hours a week.

3. Related Literature

3.1. International evidence

There is a broad range of international evidence, mainly from the US. The most famous program done is the Perry Preschool program in the United States in the 1960s. This was a good, but small, pilot to study both the short- and long run effects of ECP. A group of 123 disadvantaged African-American children were randomly assigned for either the control or experimental group. 58 children could enroll into Perry Preschool, an early intervention program. The participants were being followed until they had reached the age of 40 to study the long-term effects on earnings, employment, education, crime and a variety of other outcomes (Heckman et al, 2010). The economic effects of the Perry Preschool are striking: the benefit-cost ratio seems to be between 7 and 12 dollars. This large multiplier seems to be a result of the positive effects on socio-economic variables, such as a lower unemployment rate and lower crime rates. The effects on cognitive skills seem to fade out.

Large-scale ECP's are also studied in the US. These programs seem to have positive effects as well, but are not as strong as the small intervention programs, like Perry Preschool. Head Start is such a large-scale program in the US. This program is implemented in 50 states. Many studies are performed trying to find out the effectiveness of Head Start. Elango et al. (2015) performed a meta-analysis and reanalyzed primary data sources. They provide answers on two different questions regarding early childhood education. The first is about the effectiveness of the programs and the second is about the cost-benefit ratio. Elango et al. conclude that Head Start is effective for disadvantaged children, but the program has to meet certain quality restrictions to accomplish the effect. Early-life skills and later-life achievements improve, due to the fact that disadvantaged children spend more time in a better, stimulating environment. The effectiveness of ECP is only positive compared to home care, not if you compare it with other forms of childcare. However, studies are not unanimous in their conclusions. McKey et al (1985) did not find long-term effects on socio-economic outcomes, but a more recent study of Ludwig and Miller (2007) conclude that the program is cost-effective and therefore has a positive impact. Elango et al. discuss the way these studies are performed. They conclude that these results cannot taken too seriously, as they did not take into account facts as alternative programs children could participate in. Kline and Walters (2015) did take this into account and found negative selection into the program. After

correcting for this, they conclude that Head Start is as effective as other programs, but better than home care. Baker (2011) also mentions other difficulties in transforming the small-scale ECP's into large-scale ECP's. First, when scaling-up a program, it might be difficult to keep the quality high and keep the costs at the same level. Second, it is possible that children benefit on different levels. In the small-scale programs, extremely disadvantaged children are enrolled. These children probably benefit most of ECP. Children that are less disadvantaged, might have less effect of ECP, even when the quality is the same.

Havnes and Mogstad (2014) studied early childhood education in Norway and find that disadvantaged children benefit more from ECP than non-disadvantaged children. Children with low-income parents who were enrolled in ECP, were more likely to complete high school. Upperclass children did not profit of ECP enrollment. Besides the positive effect on high school completion, cognitive test results are not influenced by ECP. To explain the effect in educational attainment in the long-run, Havnes and Mogstad stress the importance of the development of non-cognitive skills. In the research review done by Melhuish et al. (2015) different conclusions are made: quality is very important to see any effects, mixed groups are preferable above homogeneous (disadvantaged) groups and the role of the primary school is very important. The effects of ECP will disappear when children go to a low-quality primary school. Cunha and Heckman (2007) states that the process of skill formation is dynamic and it builds on itself. ECP can be used to lay the foundation for building skills later in life. Skills are multiple, self-productive and complementary to other skills and for investment.

Considering all these studies and their conclusions, most international studies state a positive effect for disadvantaged children and stress the importance of early development. The effects of cognitive skills possibly fade-out after a few years, but effects on other aspects in life are positive. Health, education enrollment, crime rates, employment rates and wage levels seem to be positively influenced by ECP.

3.2. Evidence in the Netherlands

In Dutch politics a strong debate is going on about ECP. As earlier mentioned in the introduction, it is difficult to set up a good, reliable study in the Netherlands due to selection. And therefore studies about the effectiveness of ECP in the Netherlands need to be well evaluated. What makes it difficult to perform a good study, is a lack of good data as it does not cover a benckmark test. The methodology of a study has to take this into account to get

reliable results. Last year, two studies (Kohnstamm Instituut and Akgunduz & Heijnen) are published that both found positive effects of ECP using a reliable methodology. The Kohnstamm Instituut (2016) used *pre-COOL* cohortdata, where the aim is to get insights on the different forms of ECEC on cognitive and non-cognitive skills. Children's development is followed starting at the age of 2. It is therefore possible to get a benchmark. Unfortunately, at the start of writing this thesis, the available pre-COOL data did not have enough observations for this study.

First of all, Kohnstamm Instituut (2016) showed a clear selection in the different types of ECEC. Children that do not speak Dutch at home, with a non-Western background and where the mother is low-educated, are more often enrolled in ECP than children with Dutch as their native language and with a higher educated mother. Moreover, target children make more use of preschools than daycare, where the non-target children are equally divided between daycare and preschools. Kohnstamm Instituut studied the development of children, in the age between 2-6 years, for both ECP and non-ECP-children. They found that the implementation of ECP policy in the Netherlands is effective in stimulating the participation of target children in high quality ECP. The quality of preschools and daycare centers that offered ECP is better than preschools and daycare centers without ECP. This is also confirmed by the study of Cebeon/Regioplan (2015), who evaluated the implemented policy on education disadvantages.

Second, Kohnstamm Instituut studied the effect of ECP on cognitive and non-cognitive skills. Kohnstamm used different types of target groups, namely children with a non-Western background, children that speak a non-Dutch language at home and children with a low-educated mother. All three groups perform the same tests. They showed the average results for target and non-target children at different moments for all skills. Such that one can see the progress. The most important conclusion is that the vocabulary development of target children tends to go faster than non-target children when following ECP. This effect is smaller and not significant for children with a low-educated mother. However, for all three groups the vocabulary results were still much lower than the non-target children. For mathematics, the target children (in all three groups) score significantly lower than the non-target children, and the difference between the children is stable. Hence, no ECP effect is visible. The language skills are much lower for target children at the age of 4. However, a positive ECP effect is seen,

as the difference becomes smaller. This effect is small for the target group with low-educated mothers. The group that is selected based on the spoken language at home shows the largest effect. The main conclusion of the Kohnstamm Instituut is that ECP has positive effects, because the development of ECP children tends to go faster than non-ECP children and these effects tend to be higher in a homogeneous class. This is contradictory to the conclusion made by Melhuish et al, who plead for mixed classes.

Akgunduz and Heijnen (2016) also studied the effects of ECP in the Netherlands and use a difference-in-difference-in-difference strategy. Akgunduz & Heijnen (2016) used grade repetition as an instrument for improvement of cognitive skills. In 2012 and 2013, the 37 largest municipalities (G37) received extra subsidies to increase the quality and availability of ECP. They controlled for grade repetition in the first two years (kindergarten) for target children, for the G37 and the other municipalities who did not receive extra money. Grade repetition in second grade occurs a lot. If the teacher thinks that the pupil is not ready to proceed to the third grade, the pupil will redo second grade for another year. The extra investments in the G37 resulted in significantly less grade repetition by target children. These effects seem to be weaker in the four largest municipalities and are only seen for boys. However, extra funding for ECP is likely to have a positive effect.

Jungbluth, Nap-Kolhoff and Rodigas (2010) performed a study in Limburg where a pilot project of ECP was set up and the authors tried to quantify the effects of this pilot. They find a small, but significant effect of the program. However, mother's education and the age of the pupil explains school results better than the program itself. Moreover, Jungbluth et al state that it was far too soon to get reliable results as the program was implemented recently. Conclusions in this paper are very fragile and cannot be taken too seriously. Karssen et al. (2013) concluded that ECP-children did not have a higher level of cognitive skills than pupils who did not follow ECP, also after controlling for background characteristics. However, pupils that did not follow any form of pre-school have the lowest language skills in grade 2. This might be a sign that being involved in ECP is better than staying at home during the first 4 years (Karssen et al, 2013). Slot et al. (2013) find some small positive effects on children's development, also after controlling for selection, quality and family background. However, Slot et al. mention the low quality of ECP in most schools. This implies that we have to be careful in making conclusions about the effectiveness of the programs. Other studies conclude

that ECP is not effective in order to diminish the disadvantage of pupils. Some studies did not find any effects or only a very small effect (Bruggers et al, 2014, Fukkink et al, 2015). But these studies have their shortcomings and did not acknowledge these shortcomings. What should be considered is the fact that there is no random selection into the groups. Therefore, it is not possible to compare an experimental group with a control group. Second, ECP-children are placed for a reason. In the conducted studies they have tried to correct for this by controlling for certain background characteristics. However, it is not guaranteed that these control variables fully explain the differences. And third, there is a lot of variation between and within the groups. The intensity and duration of using ECP differs. Also the quality of the given education differs, both for ECP and non-ECP**.**

In sum, Dutch studies about the effectiveness of ECP face many difficulties to get reliable results. Recent studies show positive effects on children's development, but the majority of previous studies conclude with negative or no effects, but they are facing clear shortcomings in their study.

4. Methodology

In this paper I investigate the effect of ECP on different cognitive skills and some pupil characteristics. Cognitive skills are language and ranking skills in grade 2 and mathematical skills, vocabulary and the level of comprehensive reading in grade 5. The pupil characteristics are underachievement and popularity.

The perfect methodology would be an experiment with a treatment and control group. However, the implementation of the Dutch policy did not foresee this. Hence, it is not possible to study the effects of ECP by making use of an experimental setting. Another method to study the ECP effect would be an instrumental variable approach. A good instrumental variable is correlated with the independent variable, but should not be correlated with the error term, such that $cov(Z_i, \varepsilon_i) = 0$, with Z_i being the instrumental variable. The challenge is to find exogenous variation that affects the choice of following ECP, but not the outcome variables. However, it is most likely that background characteristics have an influence on the decision that parents make about participating in ECP, but these characteristics are also likely to influence the skill level. Hence, it is difficult to find a good instrument.

I argue that there is a negative selection effect, because the lack of random selection into ECP. Spurious correlations between ECP enrollment and unobserved family characteristics cause this negative selection. Therefore, it is not possible to compare ECP with non-ECP children. Hence, there is no instrument available to get a proper IV estimation.

Therefore, I use a different approach which makes it possible to assess the potential bias from unobserved factors. This approach is developed by Altonji, Elder and Taber (2005) and has recently been extended by Oster (2013). Both try to give an approximation of the bias that occurs if equal selection between observables and unobservables is assumed. Altonji et al. argue that the conditional independence assumption does not hold. Collecting data has its limitations. Factors that really matters and can be easily collected, are more likely to be included. Unobservable variables are left out of the dataset. Furthermore, the chosen observable control variables are likely to correlate with the dependent variable. As the control variables are not randomly selected, it is unlikely that the unobserved variables do not relate to ECP. However, this is assumed using OLS. According to Altonji et al. a selection effect is likely to occur in the decision of whether a child has to follow ECP or not. Therefore, the conditional independence assumption cannot hold. Intuitively, and also part of the Dutch

policy, children with a development disadvantage will participate into ECP. Children with a higher probability on a lower readiness to learn in grade 3 are encouraged to participate in ECP. Therefore, considering all children, the coefficient estimates of ECP, β_1, will be downward biased as these children would have a lower than average score in any case. Altonji et al. investigate the sensitivity of the estimates in case of equal selection on observable and unobservable variables. In other words, the part of an outcome that is related to the observed variables has the same relation with ECP as the part related to the unobservables. Instead of the OLS assumption that the error term ε_i of Y_i, the unobservable part, is not correlated with ECP, Altonji argues that the error term is correlated with ECP. This is a strong assumption but, according to Altonji et al, not stronger than the conditional independence assumption used with OLS. Other necessary assumptions are

1. The set of observed variables is randomly chosen from the full set of variables.
2. The number of observed and unobserved variables is large enough such that none of the elements dominates the distribution of participating in ECP.

These assumptions are not likely to hold. However, the control variables that I will include, have a broad range and probably will have substantial explanatory power for the dependent variables. But, in the steps that I will take (see Section 4.1) these control variables will lose their power in explaining ECP enrollment. The large number of control variables suggest that these are a useful guide to the unobservable variables. This suggest equal selection between observable and unobservable variables.

Before the OLS regression will be performed, a descriptive analysis will be shown in Section 5. Here, it will be clear that the ECP and non-ECP children differ in family background and characteristics. This analysis will help forming a more homogeneous subsample. This subsample includes children with more equally distributed characteristics and less differences. This way, the observable control variables will be less relevant for the dependent outcome variables. In Section 6, the ECP effect on different skills and characteristics will be estimated. In order to find the ECP effect assuming equal selection, I will perform three different steps which I will outline in the next paragraph.

4.1. Empirical approach in three steps

As already mentioned in this paper, a selection effect is likely to occur in the enrollment of ECP. The method that will be followed in this paper, will try to filter this selection effect. Therefore, three steps have to be taken.

First, I will consider the full sample, where no selection has taken place yet. I will estimate the ECP effects by performing the following two ordinary least square regressions:

$$Y_{i,g} = \beta_0 + \beta_1 * ECP + \beta_j * X_j + \varepsilon_j \quad (1)$$

$$Y_{i,5} = \beta_0 + \beta_1 * ECP + \beta_i * Y_{i,2} + \beta_j * X_j + \varepsilon_j \quad (2)$$

where Y represents the outcome variable, with i being the student and g is either grade 2 or grade 5, depending on the grade the skill level is measured in. β_0 is a constant, β_1 is the coefficient of interest and X_j implies a vector of observed control variables with β_j as their coefficients. ε_j represents the error term and its expected value is 0. Hence, it is assumed to be uncorrelated with ECP. ECP is the dummy variable that represent whether a child participated in ECP or not. For the second OLS regression, the results on cognitive skills in grade 2 are included as control variables in order to find ECP effects on cognitive skills in grade 5. By performing two different OLS regressions on the same cognitive skills, it is possible to show the importance of skill levels in grade 2. If the school results in grade 5 are depending on the skill levels in grade 2, it is likely that the early skill levels are a good predictor for future school performance. I use two different datasets for the regressions: the first dataset contains children in grade 2, and thus will estimate the cognitive skills and pupil characteristics in grade 2. For the second I use a dataset which contains grade 5 children. Test results are known from grade 2 and grade 5 and all variables will be considered that are mentioned above. I gradually will extend the group of control variables, starting without control variables. The coefficient of interest that will be estimated is $\dot{\beta}$. Next, I will add control variables on pupil characteristics, followed by an extension with family background characteristics and the parent-child interaction variables. The ECP coefficient estimated with the full set of control variables is $\tilde{\beta}$. In Section 5.1. I will explain more about the dataset and the specific (control) variables used in these regressions.

Second, I will redo the first step, but using the more homogeneous subsample. Here, $\dot{\beta}$and $\tilde{\beta}$will also be estimated. The expectation is that these estimations will be larger than in

the first step, where the full sample is used. Because the subsample is more homogeneous, the differences in school results will be smaller.

However, using these simple OLS regressions, both in step 1 and 2, will probably give biased estimates. It is likely that omitted variable bias occurs in non-experimental studies. To diminish this possibility as much as possible, control variables are included. Still, this inclusion does not cover unobservables that might be important in explaining the treatment effect.

Therefore I will do a third step in the analysis. In this step, I will, as in Altonji et al (2005), assume equal selection on observables and unobservables. This way, $\hat{\beta}$ will be estimated. Unobservable variables can be considered using the following model:

$$Y_{i,g} = \beta * ECP + \gamma_j w_j + W_2 + \epsilon. \quad (3)$$

where Y represents the outcome variable, with i being the student and g is the grade the variable is measured. β represents the coefficient of interest, ECP is the treatment and $w_j$is a vector of observed control variables, with γ_j as their true coefficient. W_2 represents a vector of unobserved control variables multiplied by their true coefficient. ϵ is orthogonal to ECP, w_j and W_2. Just as Altonji, Oster (2013) assumes equal selection. Hence,

$$\frac{\mathrm{Cov}(\gamma_1 w_1,\ ECP)}{Var(\gamma_1 w_1)} = \frac{\mathrm{Cov}(\ W_2,\ ECP)}{Var(\ W_2)}$$

The final outcome will give two estimations, $\tilde{\beta}$ and $\hat{\beta}$. $\tilde{\beta}$ will be the estimation with controls, which is already estimated in the second step. $\hat{\beta}$ represents the estimate assuming equal selection[1]. As the OLS regressions in the first two steps will probably give a downward biased result, the lower bound will be the same as the controlled estimate $\tilde{\beta}$ as is given by the OLS. The upper bound ($\hat{\beta}$) will be the bias-adjusted effect, assuming equal selection. In these steps the ECP effect will probably become larger and will come closer to zero (in case of negative coefficients). Hence, a selection effect in the enrollment of ECP is made clear.

[1] To estimate $\hat{\beta}$, two key inputs are necessary, namely δ and R_{max}. δ gives the degree of selection on observables and unobservables. Here, δ = 1, because the assumption is made that there is equal selection. R_{max} can have different values. $R_{max} = 1$ means that the outcome can be fully explained by the treatment and full control (both observables and unobservables) set. However, it is unlikely that this is the case and Oster (2013) argues that using $R_{max} = 1$, the adjustments will be too large and errors will be extremely large. According to Oster a reliable R_{max} is 1.3$\tilde{R}$, where $\tilde{R}$ is the R^2 of the controlled OLS regression.

To sum up, three steps will be undertaken. First, the ECP effect will be estimated by OLS for the full sample. Second, the first step will be replied, but using the more homogeneous subsample. Third, by assuming equal selection the unobservable variables will be taken into account and the ECP effect will again be estimated. The expectation is that using this method, the selection effect will become visible, due to the increase of the ECP effects in every step that will be taken. Thus, $\dot{\beta}$ will show the smallest ECP effect and $\hat{\beta}$ the largest ECP effect. In Section 6 the estimation results are shown and will be discussed.

5. Data

Now the methodology is considered, I continue with explaining the dataset and all variables that are involved into this study. In Section 5.4 a descriptive analysis is presented, where the differences between ECP and non-ECP children are shown. Thereafter, a subsample is formed by only considering low-SES children.

5.1. COOL – dataset and sample

The dataset used is the $COOL^{5-18}$ (from now on: COOL) cohort study, where children are followed between the age of 5 to 18, which is conducted every three year, started in 2007-2008. I used the first three cohorts for this paper. The aim of COOL is to examine the Dutch educational system by following the same pupils every cohort. The data is gathered in several questionnaires, filled in by pupils, parents and teachers, about the pupils' background, the development of their civic competences and some non-cognitive skills. Test results of the cognitive skills are made available by schools. For this paper two different groups are used. The first group consists of only grade 2 pupils, all three cohorts included. The second group includes pupils that are followed in the first two cohorts, grade 2 in the first and grade 5 in the second. Not all of the pupils that were followed in the first cohort, and could be found in grade 5 in the second cohort. Possible explanations are grade repetition, moving to another school or the school did not want to participate anymore. The researchers tried to find the lost pupils and include these in the dataset, which partly succeeded.

In total 550 schools participated in COOL. Each wave consists of a representative part (400 schools), an additional part of disadvantaged schools (130 schools) and traditional renewal schools (20 schools). The first group that will be studied in this paper consists of 21.187 pupils. Pupils in grade 2 of all three cohorts are included in this group. In the second group I included the pupils that are followed in the first and the second cohort in grade 2 and grade 5. This group is smaller, due to the sample attrition. Only 44.3% of the grade 2 pupils in the first cohorts participated in the second cohort. The researchers tried to find the 'lost' children. For a solid research it is also important to follow the drop-outs to prevent possible biases. In the end, the group consists of 8.345 pupils. From these pupils information is available about their skill level in group 2 and group 5 (Driessen et al, 2008).

Table 1
Participating pupils in COOL bij cohort and grade

	First cohort	Second cohort	Third cohort	Total	Low-SES
Grade 2	9.331	6.776	5.080	21.187	1.722
	In first and second cohort		**Moved, delayed, switch schools**		
Grade 2+5	5.877		2.468	8.345	739

5.2. Dependent and independent variables

In this section I will explain the dependent and independent variables that are used in this study. In this paper I will examine the effect of ECP on language and ranking skills in grade 2 and mathematics, vocabulary and comprehensive reading in grade 5. All these test results come from the *Cito-Leerlingvolgsysteem* (Cito-LVS). Pupils make the tests at school, and teachers grade the tests. The results are send to the COOL-researchers. The scores run from 0-100. In grade 2 pupils have to make two tests: *Taal voor Kleuters* (Language for toddlers) and *Ordenen* (Ranking). The language test consists of 56 items and it measures conceptual and metalinguistic consciousness. Conceptual consciousness is about recognizing concepts and understanding short texts. The metalinguistic consciousness is the skill to focus on the form and not the meaning of a word. It is about the focus on written language or the sound of words.

The ranking test consists of 42 items and focusses on three different aspects: the abilities to classify objects (e.g. put all animals in the same group), to rank different objects (e.g. from small to large) and the ability to compare and count objects (e.g. does this group consist of more, equal or less objects?).

In grade 5 the pupils have to make four different tests: vocabulary, the three-minutes test (which is not considered in this study), comprehensive reading and calculus/mathematics. The vocabulary test measures the vocabulary in written language. The pupils have to read a sentence, where a word is made bold. Four different meanings of the bold word are given, and the pupils have to decide which one gives the best explanation. For the comprehensive reading test, pupils have to read stories and answer questions about the story. The test consists of three different parts. The first part is based on the average skill level of pupils in grade 5. Based on the test result in this part, pupils have to make part 2 or part 3. Part 2 has a lower than average skill level and part 3 a higher than average skill level.

The last test is mathematics and consists of different aspects. Pupils are tested on their

skills on numbering, head counting, geometry, time and money.

In this paper I am also interested in pupil characteristics. In the COOL questionnaires 4 characteristics are considered, namely underachievement, behavior, popularity and work attitude. This study only considers underachievement and popularity. Pupils are scored by their teacher on a scale from 1-5, where 1 represents 'definitely not true' and 5 represents 'completely true'. For the dataset used for this paper I reconstructed the scale from 0-4. For every characteristic 3 statements are used, which have to be answered by the teacher. The statements for the characteristics are:

Underachievement:

1) This pupil's school results are representative for his/her talent.
2) The results of this pupil are lacking behind compared with his/her talent.
3) This pupil can achieve more.

Popularity:

1) This pupil can get along well with his/her classmates.
2) This pupil is popular with his/her classmates.
3) This pupil does not have many friends in class.

A low score for underachievement means that the pupil is not an underachiever, and a low score for popularity means the pupil is not very popular. A high score says that the pupil is an underachiever and is popular with his/her classmates.

In this paper I want to know the effect of ECP on the variables described above. To know whether a pupil followed ECP, I constructed a dummy variable, where ECP is 0 if the child did not enroll into ECP and 1 if the child participated. The question whether the child participated in an ECP is asked in the parents' questionnaire, but only for the children in grade 2. The dummy variable makes that coefficients are easily interpretable.

5.3. Control variables

For the regressions that will be performed in Section 6, different control variables are included. In COOL, a lot of background information is asked to the parents and teachers, to be able to control for different variables. I will control for different variables, in different steps. Table 2 shows which control variables are included in the different settings. Now, an explanation about some of the control variables will follow. Socio-economic status is based on

the education level of the parents and whether they are born in the Netherlands or not. It is divided into 6 different groups: max LBO migrant, max LBO Dutch, max MBO migrant, max MBO Dutch and HBO/WO migrant and HBO/WO Dutch. The question whether the parents are working or not is also included. The parent is considered employed if he/she works more than 12 hours a week. For the level of reading skills a question is asked where parents can answer on a scale from 1-5, with 1=not/very bad 2=bad 3=reasonable 4=good 5=really good. For the parent-child interaction variables questions are asked about how often they undertake specific activities together with their child. Questions are:

- How often do you read a book or comic together with your child?
- How often do you read out your child?
- How often are you going to the library together?
- How often do you talk about happenings at school?
- How often are you watching children programs on TV together?
- How often do you play a game (or a computer game)?

Answers run from a scale from 1-4, where 1 is every day, 2 is twice a week, 3 is a few times a month and 4 is (almost) never. For the dataset I changed the scale, from 0-3, where 0 is (almost) never and 3 is every day.

Table 2
Control variables included in different specifications of OLS-regressions

	(1) None	(2) (1) + pupil profile	(3) (2) + family characteristics	(4) (3) + interaction
Control variable				
Sex		X	X	X
Country of birth child		X	X	X
Language skills gr.2*		X	X	X
Ranking skills gr. 2*		X	X	X
Country of birth mother			X	X
Country of birth father			X	X
Socio-economic status			X	X
Family composition			X	X
Employment mother			X	X
Employment father			X	X
Language of speaking				
-between parents				X
-parent-child				X
Level of reading skills				
- mother				X
- father				X
Reading together				X
Reading out				X
Library visits				X
Daily conversations				X
Watching TV together				X
Playing games				X

* Language and ranking skills of grade 2 are included as control variables for the grade 5 cognitive skills.

5.4. Descriptive analysis of characteristics ECP and non-ECP children

To see what kind of group we are working with in this study, a descriptive analysis is performed for two separate groups. Results can be found in Table 3. Columns 1-3 are the descriptive analysis of the full sample, where ECP-children and non-ECP children are compared. Columns 4-6 are the descriptive analysis of the more homogeneous subsample of low-SES children. Columns 3 and 6 show the difference between the ECP and non-ECP children (both for the full sample and the subsample) and shows whether these differences are significant or not.

5.4.1. Full sample

First, I consider the left part of Table 3. The means of key variables are compared between children that did not follow ECP and children that followed ECP for the full sample.

The difference is compared by a t-test. Comparing these children, taking the full sample, a few things are worth mentioning. The difference between the groups is for almost every variable significant at the 1% level. This means that the non-ECP children significantly differ from the ECP-children. A simple regression about whether ECP has an effect on school results can therefore not be trusted. The family background and the environment differs between the two groups, which lead to a selection effect in the enrollment of ECP. Especially the SES of parents is lower for the ECP group. For non-ECP children, SES is significantly higher than the ECP children. Also, parents are more often unemployed and immigrants in the ECP-group.

The parent-child interaction variables are important, because these variables are likely to influence the child's development. It is expected that the more time parents spend with their children, the better they will perform. What is seen in Table 3, is that parents interact differently with their children for non-ECP and ECP children, except for reading together. However, the results are in opposite direction than expected. The parents with ECP-children more often go to the library, watch television together and play games with their children.

5.4.2. Low-SES subsample

Considering the differences between the ECP and the non-ECP group, it is likely that there is a selection effect between the two groups. Therefore, a simple comparison by performing an OLS regression is therefore not reliable. Selecting a subsample and comparing ECP with non-ECP children is therefore a more reliable method to show the effectiveness of ECP. The selection effect will decline when studying a group that is more likely to be involved in ECP. In the first three columns of Table 3 it is shown that the SES of parents significantly differs between the two groups. Therefore, I also perform a descriptive analysis on a group with only low-SES parents. Columns 4-6 show the descriptive analysis of the low-SES subsample. The total observations in this subsample is 1.722 of which 1.353 children did not follow ECP and 369 were enrolled into ECP. In column 4 and 5 of Table 3 the means are showed of the control variables, with only low-SES children included. The differences between the non-ECP and ECP-group are now smaller than in the full sample, and less significant. In the family characteristics most variables are still significantly different. What can be seen is that the ECP-group consists of more children with a non-Dutch background. Also the family composition is different between the non-ECP and ECP-group and moreover, also very different with the full sample. Mothers are in both groups more or less equal regarding employment levels. The way of interaction between the two groups is the same, only parents

with children involved in ECP, are going more often to the library. As I think this subsample is more representative than the total sample, I will perform the second step of the analysis using the subsample based on low-SES parents.

TABLE 3

Means of control variables in group 2 and group 5

	Full sample			Low-SES		
Variable	Non-ECP (N=18,283)	ECP (N=2,904)	Difference	Non-ECP (N=1,353)	ECP (N=369)	Difference
Family background:						
Female	.26	.29	-.02***	.47	.47	-0.01
Non-Dutch mother	.55	.59	-.03***	.40	.54	-.14***
Non-Dutch father	.56	.60	-.04***	.45	.58	-.13***
Socio-economic status	4.41	4.19	.22***			
Family composition	.87	.67	.20***	1.67	1.47	.21**
Siblings	.79	.78	.01	.73	.71	.02
Mother employed	.65	.57	.08***	.39	.36	.03
Father employed	.85	.78	.07***	.67	.59	.09***
Spoken language at home	.18	.32	-.14***	.40	.54	-.14***
Level of understanding Dutch, mother	3.58	3.38	.20***	3.10	2.84	.26***
Level of understanding Dutch, father	3.37	3.15	.22***	2.79	2.67	.12
Interaction parent-child:						
Reading together	3.12	3.09	.03	2.81	2.81	.00
Read out	3.06	2.90	.16***	2.68	2.72	-.04
Library	1.52	1.59	-.06***	1.52	1.64	-.12**
Having conversations	3.77	3.73	.04**	3.70	3.67	.03
Watching television	3.34	3.45	-.11***	3.49	3.48	.01
Playing games	2.90	2.99	-.09***	2.99	3	-.01
Parent involvement gr. 5[1]	3.74	3.66	.08***	3.39	3.26	.13*
Cognitive skills:						
Ranking group 2	58.91	57.12	1.79***	56.52	56.61	-.09
Language group 2	69.03	66.97	2.05***	61.00	59.87	1.13**
Vocabulary group 5	62.59	59.96	2.62***	57.58	52.03	5.55***
Comprehensive reading gr. 5	25.30	23.37	1.93***	21.09	18.88	2.21**
Maths	70.08	68.43	1.65***	66.24	64.87	1.37
Figures group 5	10.45	10.30	.16***	9.93	9.64	.29
Exclusion group 5	11.87	11.80	.07	11.31	11.19	.12
Numbers group 5	11.19	11.11	.08*	10.57	10.58	.00
Three minutes test gr. 5	71.61	72.20	-.60	71.90	71.52	.39
Category group 5	15.82	15.68	.14**	15.04	15.03	.01
Analogy group 5	13.66	13.37	.30***	12.55	12.07	.49
Non-cognitive skills:						
Work attitude gr. 5	3.42	3.38	.04*	3.32	3.25	.07
Nearness group 5	3.67	3.64	.02*	3.67	3.64	.02*
Conflict group 5	1.71	1.76	-.05***	1.72	1.90	-.18***
Independence gr. 5	2.17	2.22	-.04**	2.17	2.24	-.07
Popularity group 5	3.62	3.58	.04**	3.54	3.56	-.01
Underachiever gr. 5	2.51	2.53	-.02	2.52	2.59	-.07
Behaviour group 5	3.70	3.69	.01	3.63	3.57	.06

[1]: The variables measures in the second cohort, grade 5, have N=8.345 (N=7.279 in non-ECP, N=1.066 in ECP) in the full sample and N=739 (N=584 non-ECP, N=155 ECP) in the low-SES sample.

*, **, *** Difference is statistically significant at 10, 5 or 1 % level.

6. Main estimation results

6.1. OLS estimations

In this section I will consider the first two steps as discussed in Section 4. I will show the OLS estimates of the ECP-effect on different outcome variables, estimated with different sets of control variables, for the full and low-SES sample, using the grade 2 and the grade 2 and 5 sample. In Table 4, 5 and 6 the coefficient is shown of the ECP effect, together with the standard error and the R^2. In the first column no control variables are included, in the second column I controlled for sex and country of birth. In the third column I added the country of birth of the parents, the SES (only for the full sample), family composition, employment status, the language that is spoken at home between the parents and between parents and child and the level of reading skills. In the most extensive set of control variables I included the parent-child interaction variables. For the second group, where I test for the long-run effects, I first use the same control variables as in Table 4. These results are shown in Table 5. In Table 6 I add the control variables language and ranking skills of grade 2. The left part of the three tables consist of the OLS estimates of the full sample, the right part shows the OLS estimates of the low-SES subsample.

6.1.1. ECP effect on skills in grade 2.

First, I will focus on the complete grade 2 sample. For grade 2 two measures are available on cognitive skills, namely language and ranking skills. For the grade 2 sample the ECP effects turn out to be small, but significant. Pupils that followed ECP, score significantly lower on the language and ranking tests. The more control variables, the smaller the differences in school results. As observable variables partly explain the school results, it is a first sign of a selection effect, regarding the ECP enrollment.

As the scale on the test results runs from 1-100, the OLS results can be read as percentages. With no control variables, ECP children score 2.05 percent lower than non-ECP children for the language test. In column 4, ECP has a larger effect, as the difference in the language test becomes smaller, but it is still almost 1 percent lower than non-ECP children. For pupil characteristics, different variables are measured in the dataset. For behavior and work attitude the ECP effects were never significant. Therefore, I do not include these variables in the tables and I will not consider the results. For underachievement and popularity, following ECP turned out to have a small, but significant effect. Pupils that followed

ECP turned out to achieve more often below their skill level and are less popular among their classmates.

Second, the ECP effects are estimated for the low-SES subsample. As already expected, the ECP coefficients in the low-SES subsample differ from the coefficients in the full sample. The differences in results on language skills are still negative, but smaller and not significant in all cases. For ranking it turned out that ECP children have higher test results than non-ECP children. However, these results are not significant. The effect of ECP on underachievement is smaller in the subsample. ECP children are more often underachievers, compared with non-ECP pupils in the subsample. These results are only significant for the first two columns. The effect of ECP on popularity is very small and not significant.

Next, I perform the same OLS regression on the cognitive skills in grade 2, but with the grade 2 + 5 sample. Results can be found in Table 5. Remarkably, the ECP effects on language and ranking skills differ substantially from the results shown in Table 4. The pupils who were enrolled into ECP in the second sample had lower results for language and ranking tests than the ECP children in the first sample (only grade 2 pupils). For both samples the ECP effects on language and ranking are small and ECP children seem to score significantly lower than non-ECP children. The effect runs from 2.87 percent lower on language skills, no control variables included to 1.12 percent lower in the most extensive regression. For ranking, ECP pupils score 3.32 percent lower than non-ECP children in the regression without control variables. This effect increases to 1.72 percent in the most extensive variant. The ECP effects are less strong than in the first sample. The number of observations differs quite strongly between the two samples, as N=21.187 in the first sample and N=8.345 in the second sample. But, the pupils in the second sample are a selection of the first sample. A possible explanation is a selection effect in the follow-up of the cohort study.

In the low-SES subsample of grade 2+5 the ECP effects for language and ranking skills seems to be in line with the expectations. The differences between Table 4 and 5 are still visible, but these are largest in columns 5 and 6. For columns 7 and 8 the differences are fairly small. Effects are less negative the more control variables included. For ranking the ECP coefficient even turns positive, but is not significant. The pupil characteristics in the second subsample were all not significant. Therefore, I did not include these variables in the tables and I will not consider them.

In sum, ECP children score significantly lower than non-ECP children in grade 2. ECP

has less effect in the complete grade 2 sample, when no control variables are included. These effects increase if control variables are added. In the low-SES subsamples the ECP-effects are again higher and not always significant. This is an indication for a selection effect, considering ECP participation. If these movements also hold in grade 5 will be discussed in the next paragraph.

6.1.2. ECP effect in grade 5

Now, I perform the OLS regression on more long-term cognitive skills in grade 5. Results can be found in Table 5 and 6. In the first analysis (Table 5) I used the same control variables as in Table 4. In the second analysis (Table 6) the language and ranking skills of grade 2 are added to the control variables. Hence, if these variables partly explain the school results in grade 5, the cognitive skills in grade 2 can be an indicator for future skill levels. For grade 5 four cognitive skills are known in the dataset. Results are available on mathematics, vocabulary, comprehensive reading and technical reading (the three-minutes test). For the last variable, the results did not seem significant and are left out of this study.

Looking at the full grade 2+5 sample, the ECP effect on mathematics is only significant in the first two columns. The effect is small, ECP pupils scored 1.65 percent lower than the non-ECP children. This effect increases to a .25 percent lower score when all control variables are included. However, this result is not significant. The ECP effects on vocabulary are also negative and are spread from -2.62 (no controls) to -.87 (all controls). All these results turned out to be significant. ECP-pupils also scored lower on comprehensive reading. ECP influences the score on the comprehensive reading test with -1.93 (no control) percent to -.82 (full control) percent.

In the low-SES subsample some unexpected results are visible. It is expected that, using the subsample, the differences in results will be smaller than in the full sample. Although the ECP effects on mathematics and comprehensive reading are not always significant, it turns out that the differences in school results are larger in the subsample. ECP-children in the subsample score much lower than the ECP children in the full sample for all three variables. If we compare columns 1-4 with columns 5-8, the coefficients are smaller, more negative, in columns 5-8. For example, in the full sample ECP children score -.87 percent lower on the vocabulary test, all control variables included. In the low-SES subsample, the ECP children score 2.97 percent lower. The same holds for the test results on mathematics and comprehensive reading, but here the effects are not always significant.

In Table 6, where the language and ranking skills are added to the control variables, the ECP coefficients are higher and less significant than in Table 5. For mathematics, as soon as the extra control variables are added, the ECP effects are not significant anymore. Although not significant, the ECP effect on the mathematical skills, turns out to be positive. For vocabulary and comprehensive reading, ECP-children still have lower results, but the difference is smaller. In the low-SES subsample, the differences in coefficients between Table 5 and 6 are very small and not worth mentioning. However, a large difference in the low-SES subsample can be seen in the R^2. The R^2 is much higher if the language and ranking skills are included. Therefore, it is likely that language and ranking skills in grade 2 are a good predictor of the skills in grade 5.

In sum, ECP-children still score lower in grade 5 than non-ECP children. However, differences are not always significant. Remarkably, the ECP effects are less strong for the low-SES subsample. The differences in school results in the subsample are larger than in the full sample. Low-SES pupils that followed ECP have lower skill levels compared with ECP children in the full sample. This might be an indication of a fade-out of ECP effects, as is also seen in various international studies. Language and ranking skills in grade 2 seem to play a role in explaining grade 5 skill levels. In the full sample, the ECP effect declines when including these variables. The ECP effect in the low-SES subsample does not differ much, however, the R^2 turns out to be much higher when adding the grade 2 variables. This also might indicate an explaining role.

The OLS estimations discussed above formed the first two steps of this study. The expectations were mainly confirmed, as the full samples showed smaller ECP effects than the low-SES subsample. Only in the grade 2+5 sample the low-SES subsample showed larger differences between the ECP and non-ECP children. Because of the increase of ECP effects when including control variables and by regressing the ECP effects on a more homogeneous subsample, a selection effect is shown. However, the selection effect is not completely filtered out by this step. Therefore, I will continue with the third and final step where equal selection between observables and unobservables is assumed.

Table 4
OLS estimates of ECP effects on cognitive and non-cognitive skills in grade 2

	Full sample (N=21.187) Controls				Low-SES sample (N=1722) Controls			
	None (1)	Pupil profile (2)	(2) + family background (3)	(3) + interaction parent-child (4)	None (5)	Pupil profile (6)	(2) + family background (7)	(3) + interaction parent-child (8)
	Outcome: Language in grade 2							
ECP	-2.05***	-1.95***	-1.03***	-.91***	-1.13**	-1.02*	-.40	-.42
	(.21)	(.21)	(.20)	(.20)	(.57)	(.56)	(.55)	(.55)
R^2	.00	.06	.16	.17	.00	.02	.09	.11
	Outcome: Ranking in grade 2							
ECP	-1.79***	-1.83***	-1.04***	-.98***	.09	.11	.12	.08
	(.22)	(.22)	(.22)	(.22)	(.52)	(.52)	(.52)	(.52)
R^2	.00	.01	.04	.04	.00	.00	.03	.04
	Outcome: Underachievement in grade 2							
ECP	.03***	.03***	.01	.01	.09**	.09*	.05	.05
	(.01)	(.01)	(.01)	(.01)	(.04)	(.04)	(.04)	(.04)
R^2	.00	.01	.03	.03	.00	.02	.05	.06
	Outcome: Popularity in grade 2							
ECP	-.04***	-.04***	-.02**	-.02**	-.01	-.01	.01	.01
	(.01)	(.01)	(.01)	(.01)	(.04)	(.04)	(.04)	(.04)
R^2	.00	.00	.02	.03	.00	.01	.03	.03

*, **, ***: OLS estimate significant at resp. 10, 5 and 1% level.
Full sample consist of three cohorts of group 2, N = 21.187. The subsample consists of three cohorts of group 2 with only children of low-ses parents. The socio economic status is based on the education level of the mother. N = 1722. Control variables in column 2: sex, country of birth of child. Control variables in column 3: 2 + country of birth parents, socio-economic status, family composition and employment of parents, the language of speaking at home between mother/father and child and the level of reading skills of parents. Control variables of column 4: 3 + the level of reading together, reading out, going to the library, having daily conversations, watching TV together and playing games.

Table 5
OLS estimates of ECP effects on cognitive skills in grade 2 + 5, by sample observed in grade 2 + 5

	Full sample (N=8.345) Controls				Low-SES sample (N=739) Controls			
	None (1)	Pupil profile (2)	(2) + family background (3)	(3) + interaction parent-child (4)	None (5)	Pupil profile (6)	(2) + family background (7)	(3) + interaction parent-child (8)
	Outcome: Language in grade 2							
ECP	-2.87***	-2.79***	-1.19***	-1.12***	-2.50**	-2.08**	-.22	-.25
	(.32)	(.32)	(.31)	(.31)	(.84)	(.84)	(.82)	(.83)
R^2	.01	.03	.11	.12	.01	.04	.16	.17
	Outcome: Ranking in grade 2							
ECP	-3.23***	-3.18***	-1.79***	-1.72***	-1.18	-1.08	.19	.05
	(.42)	(.42)	(.42)	(.42)	(1.07)	(1.07)	(1.10)	(1.11)
R^2	.01	.02	.06	.06	.00	.01	.04	.05
	Outcome: Mathematics in grade 5							
ECP	-1.65***	-1.57***	-.31	-.25	-1.37	-1.57	-.47	-.64
	(.38)	(.38)	(.37)	(.37)	(1.36)	(1.35)	(1.39)	(1.40)
R^2	.00	.00	.09	.09	.00	.04	.06	.07
	Outcome: Vocabulary in grade 5							
ECP	-2.62***	-2.55***	-.92***	-.87**	-5.55***	-5.32***	-2.89**	-2.97***
	(.33)	(.33)	(.32)	(.32)	(1.17)	(1.17)	(1.15)	(1.16)
R^2	.01	.01	.18	.13	.03	.04	.15	.16
	Outcome: Comprehensive reading grade 5							
ECP	-1.93***	-1.96***	-.89**	-.82***	-2.21**	-1.86*	-1.25	-1.39
	(.33)	(.32)	(.32)	(.32)	(1.11)	(1.10)	(1.14)	(1.15)
R^2	.00	.03	.10	.11	.01	.03	.05	.05

*, **, ***: OLS estimate significant at resp. 10, 5 and 1% level.

Full sample consist of the first two cohorts: group 2 of the first cohort and group 5 of the second cohort, with the same children involved. N = 8345. The subsample consists of the same group as the full sample, but it only consists of children of low-ses parents. The socio economic status is based on the education level of the mother. N = 739.

Control variables in column 2: sex and duration of stay in NL.. Control variables in column 3: 2 + country of birth parents, socio-economic status, family composition and employment of parents, the language of speaking at home between mother/father and child and the level of reading skills of parents. Control variables of column 4: 3 + the level of reading together, reading out, going to the library, having daily conversations, watching TV together and playing games.

Table 6
OLS estimates of ECP effects on cognitive skills in grade 5, controlled for language and ranking skills in grade 2

	Full sample (N=8.345): Controls				Low-SES sample (N=739): Controls			
	None (1)	Pupil profile (2)	(2) + family background (3)	(3) + interaction parent-child (4)	None (5)	Pupil profile (6)	(2) + family background (7)	(3) + interaction parent-child (8)
	Outcome: Mathematics in grade 5							
ECP	-1.65***	.47	.20	.23	-1.37	-.68	-.51	-.61
	(.38)	(.36)	(.35)	(.35)	(1.36)	(1.23)	(1.28)	(1.28)
R^2	.00	.12	.17	.18	.00	.20	.21	.21
	Outcome: Vocabulary in grade 5							
ECP	-2.62***	-1.60***	-.55*	-.52**	-5.55***	-4.16***	-2.85***	-2.89***
	(.33)	(.32)	(.31)	(.31)	(1.17)	(1.05)	(1.06)	(1.06)
R^2	.01	.10	.18	.18	.03	.24	.29	.29
	Outcome: Comprehensive reading in grade 5							
ECP	-1.93***	-1.11***	-.52*	-.47	-2.21**	-1.21	-1.25	-1.36
	(.33)	(.31)	(.31)	(.31)	(1.11)	(1.05)	(1.09)	(1.09)
R^2	.00	.11	.16	.16	.01	.13	.14	.15

*, **, ***: OLS estimate significant at resp. 10, 5 and 1% level.
Full sample consist of the first two cohorts: group 2 of the first cohort and group 5 of the second cohort, with the same children involved. N = 8345. The subsample consists of the same group as the full sample, but it only consists of children of low-ses parents. The socio economic status is based on the education level of the mother. For the outcome variables in group 5 the language and ranking variables of group 2 are added. N = 739. Control variables in column 2: sex and duration of stay in NL, language and ranking skills group 2.. Control variables in column 3: 2 + country of birth parents, socio-economic status, family composition and employment of parents, the language of speaking at home between mother/father and child and the level of reading skills of parents. Control variables of column 4: 3 + the level of reading together, reading out, going to the library, having daily conversations, watching TV together and playing games.

6.2. Sensitivity Analysis

Now that the ECP effects are estimated by simple OLS regressions for the full sample and the low-SES sample, the results of the final step, the sensitivity analysis, will be discussed. As already mentioned, it is likely that the results showed in Table 4-6 are suffering a bias, due to omitted variables. In this section the unobservable variables will be taken into account in the regression, which will give two bounds. These bounds are given in Table 7. The lower bound will probably be the same as the OLS estimate with the full set of control variables, $\tilde{\beta}$. This is because I argue that the OLS estimates are downward biased, due to the likelihood of selection into ECP. The more disadvantaged children are enrolled into ECP, the lower school results these children are likely to have, regardless of the ECP treatment. OLS regressions are not able to clear the differences between the ECP and non-ECP children completely. Hence,

the ECP children will still score lower. The upper bound will be the ECP effect while assuming equal selection.

According to Oster (2013), there is a significant effect if zero is excluded from the bounds interval. Hence, if both bounds are either positive or negative, ECP has a significant effect on the measured skill. For each outcome different values of R_{max} are taken. R_{max} should reflect how much of the variation in the outcome variable and ECP could be explained if we had full controls for family background (all control variables). These are calculated by taken 1.3 * R^2.[2] The R^2 is taken of the full controlled OLS regressions, which can be found in Tables 4-6, column 4 and 8. Equal selection is assumed between the observable and unobservable variables, therefore δ = 1. Hence, both the observables and unobservables equally explain the treatment effect of ECP.

6.2.1. ECP effects on cognitive skills and pupil characteristics in grade 2, when assuming equal selection

In Table 7 it is shown that ECP seems to have a small, significant effect on cognitive skills and pupil characteristics in grade 2. All effects seem to be larger when the unobservable variables are taken into account. The goal of ECP is to improve language skills such that disadvantaged pupils have the same skill level as the non-disadvantaged pupils at the start of grade 3. However, when looking at the results it does not seem that this goal is reached in grade 2.

The first three columns on language skills all show negative effects. This suggests that, assuming equal selection, the results on language skills are still significantly lower for ECP children. However, we do not have a benchmark test and the fact that the differences in results are smaller than eventually thought, might be a sign that ECP has a positive effect on language skills. Furthermore, the ECP effect for the low-SES grade 2 subsample increases. The difference in the language test is very small, namely -0.2 percent. The ECP coefficients for language skills in the low-SES grade 2+5 subsample (column 4) have different signs and therefore ECP does not have a significant effect. The OLS estimate of this subsample on language skills is also not significant. Therefore, it is possible that ECP does not have any effects on language skills on low-SES children. However, looking at the grade 2 subsample with low-SES children, the bounds are significant and ECP is likely to have an effect on language skills. The difference in results between the two samples were already visible in the OLS

[2] See footnote 1, page. 17.

estimations, where I argue that a possible selection effect in the follow-up of the study might occur.

For ranking skills different results are visible. In the low-SES subsamples significant positive ECP effects are shown. Following ECP show positive results for low-SES children. The OLS estimations were not significant, but the results in Table 7 are significant and larger than the estimated OLS effects. This might be a sign that ECP has a positive effect on ranking skills, although previous studies concluded the opposite. However, we have to be careful interpreting these results as the effect for the full sample for grade 2+5 is much smaller than the effect in the full sample for grade 2. And for the low-SES subsample grade 2+5 the unobservables seem to have a much larger effect than for the low-SES subsample of grade 2. As already mentioned, this might be the result of a selection effect in the follow-up of the cohort study. Moreover, the R_{max} is also small for ranking. Hence, the control variables are not really good at explaining the ranking results. The unobservable variables that are related to the control variables are therefore not very informative either. Hence, it is not unlikely that using other control variables with a better fit, show other results.

The results on pupil characteristics underachievement and popularity do not differ much. The interval between the bounds and R_{max} are very small, suggesting that the unobservable variables do not play a role in the effects.

6.2.2. ECP effects on cognitive skills in grade 5 when assuming equal selection

When looking to the more long-term ECP effects, we see a large difference between the lower and upper bound for mathematics in the grade 2+5 sample, where the skill levels of grade 2 are included as control variables (columns 5). The OLS ECP effect was already positive, but assuming equal selection, the effect of following ECP turns out to be very positive. ECP children almost score 1 percent higher than children that did not follow ECP. This result does not hold for the low-SES pupils (column 6). Low-SES children that follow ECP have lower test results on mathematics than the low-SES children that did not follow ECP. This can be a sign that ECP is useful for all children and not only for target children. This is a finding the OECD (2016) supports. One of the main conclusions in the OECD review was that the Dutch educational system can be improved by investing in ECP for all children.

Table 7
Estimates of the ECP effect assuming equal selection on observables and unobservables

	Gr. 2 (1)	Gr. 2 low-ses (2)	Gr. 2+5 (3)	Gr. 2, 5 low-ses (4)	Gr. 2+ 5 (5)*	Gr. 2+5 low-ses (6)*
	A. Language group 2					
R_{max}	0.22	0.14	0.16	0.22		
	(**-0.90**, -0.57)	(**-0.42**, -0.20)	(**-1.14**, -0.49)	(**-0.25**, 0.53)		
	B. Ranking group 2					
R_{max}	0.05	0.055	0.08	0.06		
	(**-0.97**, -0.77)	(**0.075**, 0.08)	(**-1.75**, -1.28)	(**0.05**, 0.34)		
	C. Underachievement					
R_{max}	0.04	0.08				
	(**0.003**, 0.009)	(**0.04**, 0.05)				
	D. Popularity					
R_{max}	0.03	0.04				
	(**-0.02**, -0.02)	(**0.01**, 0.02)				
	E. Mathematics					
R_{max}			0.12	0.08	0.23	0.28
			(**-0.24**, 0.17)	(**-0.64**, -0.47)	(**0.13**, 0.96)	(**-0.61**, -0.37)
	F. Vocabulary					
R_{max}			0.17	0.21	0.24	0.38
			(**-0.87**, -0.29)	(**-2.97**, -1.95)	(**-0.65**, 0.65)	(**-2.89**, -1.96)
	G. Comprehensive reading					
R_{max}			0.14	0.07	0.21	0.19
			(**-0.82**, -0.46)	(**-1.39**, -1.10)	(**-0.60**, 0.27)	(**-1.36**, -1.09)

*: The control variables language and ranking skills are included.
Bounds are taken for the full controlled set of variables: sex, country of birth of child, country of birth parents, socio-economic status, family composition, employment of parents, the language of speaking at home between mother/father and child, the level of reading skills of parents, the level of reading together, reading out, going to the library, having daily conversations, watching TV together and playing games. Bold numbers are the controlled estimates ($\tilde{\beta}$) as shown in Tables 4-6.

Stimulating the development of young children by offering them ECP can improve their cognitive skills and eventually, it can upgrade the Dutch skill level in general.

For vocabulary and comprehensive reading, the ECP does not seem to be significant when controlling for ranking and language skills. However, the upper bound is positive. Therefore, assuming equal selection, ECP might have a positive effect. For the low-SES group the effect is smaller and differences in test results are large. Assuming equal selection makes the effect somewhat less negative. Just like mathematics, the same conclusion can be drawn for these two variables.

When excluding the ranking and language skills in the control variables (columns 3-4), ECP seems to have less effect on vocabulary and comprehensive reading in the full sample. It is possible that the language and ranking skills in grade 2 are of such importance in explaining

future cognitive skills that the significance of ECP disappears at the moment we control for these cognitive skills. Due to the fact that the effect is not significant after assuming equal selection, the conclusion that the unobservable variables play a significant role in explaining the effects on vocabulary can be made. Therefore, we should not rely too much on the negative effect found in the OLS estimation.

6.3. Summary of main results

In the last two paragraphs, three steps are taken to show the selection effect regarding ECP and the difference in ECP effects on various variables that occurs due to this selection effect. What is shown is that by each step that is taken, the ECP effects increased and differences in cognitive skills and pupil characteristics declined. Hence, by each step, the selection was partly filtered out. The OLS estimations on the full samples gave a downward biased effect. When adding control variables, the ECP effects increased. Moreover, taking the subsamples of low-SES pupils, the majority of the ECP effects increased even more. By using equal selection on observables and unobservables, the effects of ECP are less negative or even positive. Especially for the low-SES pupils, the difference between ECP and non-ECP children in language skills in grade 2 is negligible. For ranking, the ECP children perform even better. This also holds for the pupil characteristics underachievement and popularity. On the long-run, ECP does seem to have less effect on the children. The effect is either not significant or the differences between ECP and non-ECP children are still quite large. However, concluding that ECP has a negative effect or is ineffective is made too soon. There is no guarantee that all selection is filtered out by these steps. Differences in school results might be even smaller, in case of more selection. Also, by lack of a benckmark test, it is not possible to say something about the marginal development of these children.

7. Discussion

7.1. Main results compared with existing literature

Unfortunately, the results presented in this paper are not unambiguous. The different samples show different effects. The expectation was that the estimations in the low-SES subsamples would be less negative or more positive than in the full samples. Moreover, the ECP effect will increase even more when equal selection is assumed. This is only partly true. In the short-run the ECP effects are indeed less negative or even positive if we compare the full sample and the low-SES subsample. However, in the long-run the ECP effects are much smaller for the low-SES subsample than the full sample. How do these results compare with the previous findings from literature?

The effect on vocabulary in this paper is quite small for the low-SES subsamples. The Kohnstamm Instituut concluded that ECP-children developed their vocabulary faster than non-ECP children, but the ECP-children still scored significantly lower than the non-ECP children. In this paper, it is not possible to say something about the marginal effects of development, but the differences in vocabulary are still quite large when comparing ECP with non-ECP pupils. Therefore, these results seem comparable.

For mathematics the Kohnstamm Instituut did not find any significant ECP effect. This coincides partly with the results found in this study. In the full sample without the control variables language and ranking skills in second grade, no significant effect is found. But in the full sample and considering the grade 2 variables, positive effects of ECP are found on mathematics. In the low-SES subsamples ECP children have lower results than non-ECP children.

The ECP children score significantly lower on language skills, according to the Kohnstamm Instituut. Results in this paper only found small differences in the test score. Hence, ECP seems to have a larger effect in this study than Kohnstamm suggested.

An explanation for the large differences in the long-run for the low-SES subsamples can be that the ECP effect fades out. In this case, the results coincide with international evidence, where cognitive effects seem to fade out. However, in these international studies, a positive effect eventually occurs, especially when considering aspects as employment levels, wages and crime rates. The COOL data is not sufficient enough to measure these very long-term effects. It would be interesting to study these effects in the future.

Overall, the ECP effects found in this study are comparable with previous Dutch studies. ECP children still have lower school results than non-ECP children. However, what is shown in this study and that is not taken into account in most studies, is the selection effect that has taken place. Therefore, the conclusion of this paper is different than most previous Dutch studies. As a large selection bias into ECP is shown, it is expected that ECP children have lower school results than non-ECP children. When filtering out this selection, the ECP effects increase. Hence, assuming that not all the selection is taken away by performing this study, it can be concluded that ECP is not as ineffective as previously thought. There is a reliable chance on positive effects, as Kohnstamm suggest. Hence, a simple conclusion that ECP does not have a positive effect on cognitive skills is made too easily.

7.2. Limitations and recommendations

Now all the results are explained and discussed, I will mention a few limitations of this study, which have to be taken into account. Moreover, a few recommendations will be given for future research.

First of all, the dataset was the best one available, but it had its shortcomings. There was no benchmark test available, because the gathering of the data started in grade 2, after the ECP participation. The only question which was related to ECP enrollment, was the question: Did your child participate in ECP? Therefore, it would have been better to work with the pre-COOL cohort dataset, which is the follow-up of COOL. Unfortunately, at the time of writing the available data had limited observations for ECP-children in grade 2 and 4[3]. Therefore, the results would not be reliable. In the nearby future the dataset has a sufficient amount of observations. Hence, I strongly advice to study the effectiveness of ECP on different skills using pre-COOL data.

Second, the second cohort suffered from sample attrition. It is tried by the researchers to find the 'lost' children, which partly succeeded. However, the language and ranking results in the grade 2+5 sample strongly differ with the grade 2 sample. Hence, there is a chance a bias occurred. If selection indeed occurred, there is a chance of biased effects on the long-term cognitive skills.

Third, some variables suffered a lot from missing values. Not all schools are testing their pupils in grade 2 on language and ranking for example. In this study, these missing values

[3] The Pre-COOL dataset measures children's skills every two years instead of three years in the COOL dataset.

are conducted into the mean of the sample, but it is possible that the results are biased this way.

Fourth, in this study the quality of ECP is not taken into account. Previous Dutch studies mentioned the low-quality of ECP, before the implementation of the OKE-act in 2010. From other studies it is known that good quality is a strict condition to enhance skills. As the dataset has observations of 2007 and 2010, it can be expected that the quality was not as high as it should be. In this case, low quality might be the cause of observed negative effects.

Fifth, the subsample of low-SES pupils is one example of a more homogeneous group. More homogeneous subsamples are possible, like is done in the Kohnstamm Instituut study. They used three different subsamples: children with a non-Western background, a low-educated mother and children that do not speak Dutch at home. These three subsamples showed different ECP effects. This study can be repeated with different subsamples as well. It might be interesting to see if different effects are shown.

The last limitation is the fact that the only comparison made is ECP children with non-ECP children. It might be interesting to divide the non-ECP children into separate groups: children that went to daycare, children that went to pre-school and children that stayed at home. As is concluded in previous research, children of low-SES parents that stay at home before primary school have the lowest results in primary school. Comparing the cognitive and non-cognitive skills between these groups might give interesting results. I strongly advice to investigate the ECP effects on various skills compared with these different groups.

8. Conclusion

This study assessed the effectiveness of ECP on cognitive skills and pupil characteristics, while assuming equal selection on observable and unobservable variables. Daycare centers and pre-schools could start with special early childhood programs in 1998. Main goal was to improve the language skills of young children, and better prepare pupils to start learning in grade 3. This could be reached in a playful manner, for example by doing language games and reading out. In 2010 the government put a higher focus on ECP, by implementing the OKE-act. The start of this act meant higher requirements, more investment and a stricter policy for ECP. This new policy was a result of international evidence on the importance and effectiveness of ECP.

However, Dutch studies could not find the same positive results about ECP in the Netherlands. This is caused by the lack of a good study design and negative selection that is likely to occur in the enrollment of ECP. A possible study design is the methodology Altonji et al. (2005) developed and was further extended by Oster (2013). This method assumes equal selection on observable and unobservable variables, and can show the selection that has taken place. This paper uses this methodology with the COOL cohort-dataset, which covers a lot of information and background information about pupils aged 5 to 18. Three steps had to be taken to get the final results. First, ECP effects are assessed by using OLS, using the full sample. Control variables are gradually added to the regression. Hence, $\dot{\beta}$ is estimated and covers the ECP effect without control variables. Next, $\tilde{\beta}$ is the ECP effect with all control variables included. Second, the same regressions are performed, but with a subsample of low-SES children. ECP and non-ECP children are compared, but due to the more homogeneity in the sample, the ECP effects should be larger compared with the effects found in the first step. Third, equal selection between observables and unobservables is assumed. Here, $\hat{\beta}$ is regressed. It is expected that $\hat{\beta}$ is smaller than $\tilde{\beta}$. Hence, $\dot{\beta} < \tilde{\beta} < \hat{\beta}$.

On the short-run, it turned out that ECP effects are smaller when the full sample is considered. Low-SES children are more homogeneous and show less differences between ECP and non-ECP children. However, this result did not hold in grade 5. The differences in school results in grade 5 are larger for low-SES children than in the full sample. A selection effect in the second cohort of COOL might be a reason for these results. Moreover, international evidence (Heckman et al, 2010) states a fade-out of cognitive skills in the medium-run. This is

also a possible explanation. However, the expectation that the ECP effect becomes less negative or more positive when adding control variables holds for all estimated results. Hence, $\dot{\beta} < \tilde{\beta}$. In the last step, all regressions are redone, but assuming equal selection. The expectation is confirmed here as well, $\tilde{\beta} < \hat{\beta}$. The negative ECP effect gradually turns to zero when filtering the selection effect. Hence, it can be concluded that negative selection has taken place.

Conclusions about the effectiveness of ECP are difficult to make. Filtering out the selection is tried, but there is no guarantee that all selection is excluded by this method. If more selection has taken place, it could be that, in the end, no difference in school results occurs between ECP and non-ECP children after following ECP. And, as Kohnstamm Instituut (2016) and Akgunduz & Heijnen (2016) have studied, ECP children develop in a faster way than non-ECP children. This could not be verified by this study, but it is an important finding in Dutch literature.

Considering all these results and the actual debate in the Netherlands, further research is necessary. It is important to study the ECP effects for a longer period, with better data. International evidence stresses the importance of high quality ECP and that preventing a development delay is better than fighting a development delay. A clear conclusion does not arise from this study. What can be said is that ECP is not as ineffective as a lot of Dutch studies suggest. By assuming equal selection on both observable and unobservable variables, the effects turn out to be less negative as eventually thought by some researchers. Sometimes effects even turn positive. Moreover, assuming that the quality of ECP before 2010 was not very outstanding, I would say that ECP can be effective in preventing development delay when it meets certain requirements. Also having the recent results of Kohnstamm Instituut (2016) and Akgunduz & Heijnen (2016) in mind, I would say that ECP deserves more time to convince others of its effectiveness. Hence, further research is necessary.

Bibliography

Akgunduz, Y. E. & Heijnen, S. (2016). *Impact of funding targeted pre-school interventions on school readiness: Evidence from the Netherlands. (CPB Disucssion Paper 328).*

Altonji, J.G., Elder, T.E. & Taber, C.R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economics,* 113-1, pp 151-183.

Anderson, L.M., Shinn, C., Fullilove, M.T., Scrimshaw, S.C., Fielding, J.E., Norman, J. & Carande-Kulis, V.G. (2003). The effectiveness of early childhood development programs: a systematic review. *American Journal of Preventive Medicine*, 24-3S, p. 32-46.

Baker, M. (2011). Innis Lecture: Universal early childhood interventions: what is the evidence base? *Canadian Journal of Economics,* 44(4): 1069-1105.

Brandpunt (2015, Nov. 3). Retrieved from: http://www.npo.nl/brandpunt/03-11-2015/KN_1675263

Camilli, G., Vargas, S., Ryan, S. & Barnett, W.S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112-3, p. 579-620.

Centraal Bureau voor de Statistiek (2015, Sep. 22), *Peuters lage-inkomensgroepen blijven vaker thuis* [Press release]. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2015/39/peuters-lage-inkomensgroepen-blijven-vaker-thuis

Cebeon/Regioplan (2015). *Evaluatie specifieke uitkering en gemeentelijk beleid inzake onderwijsachterstanden: Gemeentelijke bestedingen en beleid 2012-2014. Voor- en vroegschoolse educatie.* Retrieved from: http://www.cebeon.nl/wp-content/uploads/2015/04/14827-Evaluatie-SU-OAB.pdf

CPB (2016). *Kansrijk onderwijsbeleid*. Den Haag: Centraal planbureau.

Driessen, G.W.J.M. (2004). A large-scale longitudinal study of the utilization and effects of early childhood education and care in the Netherlands, *Early Child Development and Care*, 174(7-8), p. 667-689.

Driessen, G. & Doesburgh, J. (2003). *Voor- en vroegschoolse educatie en cognitieve en niet-cognitieve competenties van jonge kinderen.* Nijmegen: ITS.

Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J. & Veen, I. van der. (2009). *Cohortonderzoek $COOL^{5-18}$. Technisch rapport basisonderwijs, eerste meting 2007/08*. Nijmegen: ITS/Amsterdam: SCO-Kohnstamm Instituut.

Elango, S., Garcia, J.L., Heckman, J.J. & Hojman, A. (2015). *Early Childhood Education*. (NBER Working Paper, No. 21766).

Fukkink, R., Jilink, L. & Oostdam, R. (2015). *Met een blik op de toekomst: Een meta-analyse van de effecten van vve op de ontwikkeling van kinderen in Nederland.* Kenniscentrum Onderwijs en Opvoeding Hogeschool van Amsterdam.

Havnes, T. & Mogstad, M. (2014). *Is Universal Child Care Leveling the Playing Field?* (Discussion Paper, No. 774), Statistics Norway, Research Department.

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. & Yavitz, A. (2010) The rate of return to the high/scope perry preschool program. *Journal of Public Economics*, 94-1:2. P.114-128.

Inspectie van het Onderwijs (2013). *Extra aandacht nodig voor achterstanden bij het jonge kind: Eindrapport bestandsopname voor- en vroegschoolse educatie in Nederland.*

Inspectie van het Onderwijs (2014). Tussenrapportage VVE in de G37 monitor kwaliteit van voor- en vroegschoolse educatie in de 37 grote steden in 2013 en 2014. Retrieved from: http://www.onderwijsinspectie.nl/binaries/content/assets/publicaties/2014/tussenrapportage-vve-in-de-g37.pdf

Jungbluth, P., Nap-Kolhoff, E. & Rodigas, E. (2011). *Peuters en Kleuters in het Zuid-Limburgse Onderwijs. De (voorlopige) evaluatie van de Zuid-Limburgse VVE-pilot 'Moelejaan' 2009-2010.* Universiteit Maastricht.

Karssen, M., Veen, I. van der, Veen, A., Daalen, M. van, & Roeleveld, J. (2013*). Effecten van deelname aan en kwaliteit van voor- en vroegschoolse educatie op de ontwikkeling van kinderen.* Rapport 894, Amsterdam: Kohnstamm Instituut.

Kline, P. & Walters, C. (2015). *Evaluating public programs with close substitutes: The case of Head Start.* (NBER Working paper, No. 21658).

Kohnstamm Instituut (2016). *Pre-COOL cohortonderzoek. Effecten van kwaliteit van voorschoolse instellingen.* (Rapport 947, projectnummer 20696). Amsterdam.

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(1): 159-208.

McKey, R.H., Aitkan, S.S. & Smith, A.N. (1985). *The impact of head start on children, families and communities.* (United States Head Start Bureau). Washington, DC.

Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K. … & Leseman, P. (2015). *A review of research on the effects of early childhood education and care on child development.* Utrecht University Repository.

Ministerie van Onderwijs, Cultuur en Wetenschap (2013). *Kwaliteit van voor- en vroegschoolse educatie.* Retrieved from: https://www.rijksoverheid.nl/documenten/kamerstukken/2013/08/21/kwaliteit-van-voor-en-vroegschoolse-educatie

NRC Handelsblad (2016, June 7). *Nu blijkt ineens: taalles voor jonge kinderen werkt dus wel*. Retrieved from: https://www.nrc.nl/nieuws/2016/06/07/nu-blijkt-ineens-taalles-voor-jonge-kinderen-werk-1626733-a1508642

OECD (2015), *Education at a Glance 2015: OECD Indicators,* OECD Publishing. Paris. Retrieved from: *http://dx.doi.org/10.1787/eag-2015-en*, p. 333, Table C2.1.

OECD (2016), *Netherlands 2016: Foundations for the Future*. OECD Publishing, Paris.

Paas, T., Mulder, L. & Roeleveld, J. (2013). *Zittenblijvers en verwezen leerlingen in het cohortonderzoek* $COOL^{5-18}$. Nijmegen: ITS/Amsterdam: Kohnstamm Instituut.

Veen, A., Veen, I. van der, Karssen, A.M. & Roeleveld, J. (2013). *Deelname aan voor- en vroegschoolse educatie en de ontwikkeling van kinderen. Twee onderzoeken in het kader van de BOPO-themalijn 'effecten van beleidsontwikkelingen in het onderwijsachterstandenbeleid'. (*Rapport 901, projectnummer 20442 & 20447). Amsterdam: Kohnstamm Instituut.