Stockholm School of Economics
Department of Economics
5350 Master's thesis in economics
Academic year 2017–2018

The Effect of Relative School Starting Age on Having an Individualized Curriculum in Finland

Aapo Kivinen (41062)

Abstract: In Finland, a child in special education receives an individualized curriculum when standard support does not suffice. One factor that may have an impact on the assignment of an individualized curriculum is the relative age of the child. Due to the cutoff date of school starting age, there is an age gap of roughly one year in each class. This difference in relative age can affect through few possible mechanisms: difference in absolute age, peer effects, and the optimal school starting age. In this thesis, I study how relative school starting age affects the probability of having an individualized curriculum.

I use regression discontinuity design and individual level register data for middle school graduates in 1998–2014 to estimate the causal effect of relative school starting age. Relatively younger graduates are 1.4 percentage points more likely to have a partially individualized curriculum than graduates who are a year older. Respectively, older graduates are 1.8 percentage points more likely to have a regular curriculum. The results are robust and they hold for multiple specifications. I also find that the relative age effect is stronger for girls and students with lower educational background. Furthermore, when studying temporal variation of the effect, I observe a significant effect only from 2005 onwards. This may be partly explained by the curriculum reform in 2004. My research contributes to the areas of special education and relative age effect. The results are in line with prior literature of relative age.

Keywords: Special Education, Relative Age Effect
JEL: I20, I21, I28

Supervisor: Anders Olofsgård
Date submitted: December 1, 2017
Date examined: January 9, 2018
Discussant: Malte Joost Truelsen
Examiner: Maria Perrotta Berlin

Acknowledgements

I want to express my gratitude to Antti Saastamoinen at VATT Institute for Economics Research for guiding me and always offering valuable comments while writing this thesis. I also want to thank Mika Kortelainen, Martti Kaila, and Kristiina Huttunen at VATT for the advice and comments.

Special thanks to Anders Olofsgård, my supervisor at SSE. Discussions with him were encouraging and helped to clarify my thinking. Furthermore, I want to thank Markku Jahnukainen at the University of Helsinki and Tanja Kirvainen at the National Audit Office of Finland for offering me precious comments and important insights from the field of special education.

I want to thank Rachael for the review she provided during the process. Lastly, I want to acknowledge my dearest Hanna for being patient and supportive along this project. Thanks for your endless comments and discussions.

Contents

1. Introduction

According to Finnish Basic Education Act (628/1998), every child has an equal right for support and guidance during their basic education. A common way of support is special education which is divided into three tiers based on the need of the support. There are many factors that can affect the child's need of support. Family background, personal capabilities and classmates are just few examples of these factors (Kirjavainen et al., 2014). There is also evidence that relatively younger children have worse educational outcomes and are more likely to receive special education (Dhuey and Lipscomb, 2010; Fredriksson and Öckert, 2014; Kaila, 2017). In literature, this effect of relative school starting age is called *the relative age effect*. In this thesis, I study the relative age effect on individualized curriculum, which is another way of support for special education students.

My first research question is how relative school starting age affects the likelihood of middle school graduates to have an individualized curriculum. In Finland, a special education student has Individual Educational Plan which depends on the severity of the disability. The plan states whether the student follows regular curriculum, an individualized curriculum in some subjects or in all subjects (Kirjavainen et al., 2016). A little more than half of the special education students have at least one individualized subject[1]. In the data, I have information on the level of individualization of each student but I cannot separate special education students with regular curriculum from students in standard education. Thus, I use the level of individualization as an outcome of interest rather than special education status.

My second research question is whether the relative age effect differs for middle school graduates by gender or parents' educational background. Research has shown that demographic background, including socio-economic status and mother's education level, affects the probability of receiving special education (Kirjavainen et al., 2014). Thus, it is relevant to study whether these factors

[1] For details about the numbers of special education students and students with individualization, see Table A.1 in Appendix.

affect the relative age effects.

Special education is a politically relevant topic and it has been reformed multiple times in Finland in the 2000s (Thuneberg et al., 2014). Through some of these reforms, special education has become more inclusive and its resources have increased (Kirjavainen et al., 2016; National Audit Office of Finland, 2013). Kirjavainen et al. (2016) state that these reforms may also be a reason for the increase in the number of special education students during the last twenty years. My third and final research question is whether the relative age effect in individualizing children's curricula has changed over time.

There are two different categories of individualization: *partially individualized curriculum* (PIC) and *mainly or fully individualized curriculum* (MOFIC). MOFICs are mostly assigned to children with severe disabilities and I do not expect relative age to have an effect on their diagnosis. In contrast, PIC is often assigned to children who have difficulties with a certain subject and, thus, relative school starting age can have an impact on them.

As a result of this thesis, I find that one year in relative age decreases the probability of having PIC by 1.4 percentage points for middle school graduates. In relative terms, it is a large increase in probability since only 2.8% of all graduates have a PIC. Similarly, I estimate that a year older graduates are 1.8 percentage points more likely to have a regular curriculum. However, I do not find a robust, statistically significant effect for graduates with MOFIC. As a result for the second research question, I find that the relative age effect is larger for females in PIC. The effect is also stronger for graduates whose parents have a lower educational background in both regular curriculum and PIC.

The relative age effect has varied over time. Until 2005, the relative age effect of one year has not been statistically significant in the probability of having a PIC. From 2005 onward, I find a statistically significant effect that students who are a year older in school starting age are 2.1 percentage points less likely to graduate with a PIC from middle school.

In my identification strategy, I exploit the school starting age rule, which assigns children to start school in August during the year they turn seven (Finnish Basic Education Act, 628/1998). Therefore, children who are born before and after New Year are almost same age but the ones who are born after New Year start the school a year later than the ones who are born before.[2] However, the school starting age rule is not binding so some children may start school a year earlier or later than assigned to. To overcome this challenge, I use children's birth date to estimate the probability of school starting age. This method is called fuzzy regression design (Imbens and Lemieux, 2008). Due to this identification strategy, I can interpret my results as a causal effect of the relative school starting age on the outcome. My results are internally valid but I cannot expect the results to hold externally, for example, in other countries or extrapolating the relative age difference to more than a year.

It is important to study the relative age effect since it is not definite whether relatively younger children need more special education or if their need for special education is misdiagnosed. Students, particularly with moderate difficulties, may suffer from the stigma of special education through lower expectations from family and teachers and lack of effort (Keslair et al., 2012). However, it could also be the case that younger students benefit from supplementary support since they are in a disadvantaged position compared to their older peers.

There are few possible mechanisms for how school starting age affects outcomes and some of them are difficult to distinguish. First of all, older students start school at an older age meaning that they have accumulated more skills before starting school (Elder and Lubotsky, 2009). They are also older at every moment of evaluation and exam, which gives them an advantage. Another possible mechanism is the immaturity of younger students compared with their peers. Teachers can interpret this immaturity as lower ability and, hence, be more likely to assign younger students to special education (Dhuey and Lipscomb, 2010). These above-mentioned mechanisms give an advantage for

[2] In this thesis, I use the term New Year frequently, so it is important to define it precisely. By New Year I mean the change of year and it does not refer to any specific date. So, December 31 is still before New Year and January 1 is after New Year.

older students. Moreover, relative age can affect students through peer effects. Unlike previous reasons, these can affect the outcomes to both directions. Being among the youngest in class can motivate to work harder and catch up with others or it can lead to lower confidence and giving up. Even if we cannot distinguish between different causes of relative age effect it is important to discuss them. Different mechanisms would lead to different policy advice and, thus, further research for these is required.

My thesis builds on two research areas: relative age effect and special education. Despite the rich literature on relative age effect, there is not much research on its impact on special education. I am only aware of the study conducted by Dhuey and Lipscomb (2010). They study how school starting age affects the special education status of students in the United States. The authors use a representative sample and linear instrumental variable estimation. I study this topic in a Finnish context. I use detailed register data and a fuzzy regression discontinuity design. Furthermore, there is not much causal literature on the assignment to special education in Finland. My thesis contributes by providing understanding on how relative school starting age affects the likelihood of having an individualized curriculum.

The thesis proceeds as follows: In section 2, I introduce the relevant background for the thesis. I explain the Finnish basic and special education system before continuing to special education research both in Finland and elsewhere. I complete the section 2 by presenting the definition of relative age effect and summarizing the relevant research on relative age effect. After that, in section 3, I move on to explain the methodology I use for the estimations. I also present the main assumptions of my methods. In section 4, I present the data and the restrictions I make to my sample. I start section 5 by discussing the validity of my setting. Then, I present my empirical work, in both graphical and numerical forms. I do several robustness analyses for my estimations. In section 6, I discuss the results, their implications, and limitations. Finally, I conclude the thesis in section 7.

2. Background

2.1. Basic and Special Education in Finland

Basic education in Finland is regulated by the Finnish Basic Education Act (628/1998) and its amendments. It sets the objectives of basic education, to whom basic education is applied to, and how it is organized. It also sets the framework for the national curriculum and regulates pre-primary education. More detailed regulations, objectives, and pedagogical means are determined in the Core Curriculum which is governed by the Finnish National Agency for Education[3]. The Core Curriculum has been reformed every ten years and the latest reforms for basic education have occurred in 2004 and 2014. (Lahtinen and Lankinen, 2015, p. 124–125).

Children in Finland start their school in August during the year they turn seven. Compulsory education lasts until students have finished their nine years of basic education (6 years of primary school and 3 years of middle school) or ten years after starting school, which for most students is the summer of the year they turn seventeen. One implication of this system is that children who are born around New Year are close to each other in absolute age but very different in relative age compared to their own cohort.

Basic education has two main objectives: to support children's growth and to educate basic skills and knowledge that are needed in life. According to the Finnish Basic Education Act (628/1998), basic education should also be "*promoting civilization and equality in society and pupil's prerequisites for participating in education and otherwise developing themselves during their lives.*" The act states that education should secure regional equity throughout the country. As a part of the foundation of education, the act declares that "*education shall be provided according to the pupil's age and capabilities* - ." This statement provides the basis for special education. Since education is provided

[3]Before the merge with the Center for International Mobility in 2017, the bureau was called the National Board of Education.

flexibly according to pupil's capabilities, it means that less advantaged pupils should be provided special help. The law itself provides a variety of possibilities to offer individualized arrangements of education. For example, the basic education law allows children to begin school a year earlier or later than usual, provides rules for educating immigrant children, and guides the arrangement of special education. (Lahtinen and Lankinen, 2015, p. 225–226).

There were multiple reforms in special education in Finland in the 2000s. These reforms may have affected the trends in special education in both directions. For example, LATU project,[4]which launched by the National Board of Education in 2002, was designed to improve the capabilities of the municipalities to support students with special needs. The project was based on an earlier project called Laatu 1997–2001[5], which aimed to provide research information regarding the state of special education in Finland (Naukkarinen, 2005). The motivation of the research project was to bring the idea of inclusion and integration of special education students to general education (Naukkarinen, 2005). Based on these ideas the LATU project made special education more organized on a local level and improved the cooperation between different operators. This type of project may have increased the number of students in special education due to raised awareness and an increase in supply. (Finnish National Agency for Education, 2017).

Reforms in the Core Curriculum have also had an impact on the developments in special education. One of the objectives of the Core Curriculum Reform in 2004 was to "*emphasize student's well-being and prevent social exclusion.*" The reform wanted to "*to take individual needs into account better than earlier and to build a system for students who need special support inside the basic education system.*" The reform led to an emphasis on early detection of learning difficulties, early provision of the needed support, and the importance of multi-professional cooperation. These developments may have been a cause of an increase in milder disabilities in special education. (Finnish National

[4]Erityistä tukea tarvitsevien opetuksen kehittäminen yleis- ja erityisopetuksessa 2002–2004 [LATU – The development of special needs education in general and special education 2002–2004].

[5]Erityisopetuksen laadullinen kehittäminen 1997–2001 [The qualitative development of special education 1997–2001].

Agency for Education, 2007).

In 2009, the funding of special education was reformed from enrollment based to capitation. Before the funding reform, the government compensated municipalities on special education based on the number of students enrolled in the special education program. With this system, municipalities with weaker financial situation were more likely to enroll students in special education in order to receive more resources (National Audit Office of Finland, 2013). To cease the increase of special education students, the government reformed the funding to be based on the overall number of residents in the age of compulsory schooling (6 to 15) in the municipality. As Dhuey and Lipscomb (2013) state, governments can save resources when the funding is based on capitation. The authors call this type of arrangement census funding, meaning that the funding which regions receive is based on their population characteristics. However, this type of funding neglects some special, unobserved, features of the regions. This neglect may lead to unequal funding of different districts. The researchers suggest that some of these inequalities can be lessened by taking into account, for example, socio-economic or racial differences between regions.

In 2010, the Basic Education Act reform was introduced. The reform was developed along with the KELPO initiative[6], which was the largest implementation of a reform in Finnish basic education in the 2000s (Ahtiainen, 2010). The aim of the initiative was to implement the special education strategy of the Finnish Ministry of Education (see Ministry of Education, 2007). The strategy included suggestions about early intervention, a model of inclusive special support, and the introduction of intensified support. The KELPO initiative was a multi-phase project, starting from the strategy and development leading to the introduction of the act and implementation of the reform. In 2008, 233 municipalities out of 415 were in the initiative and by 2011, the number was 270 out 336[7]. In the same period, the government transferred over 45 million euros for the development of special education in municipalities. (Rinkinen and Lindberg, 2014).

[6] Tehostetun ja erityisen tuen kehittämistoiminta [The development of intensified and special support].
[7] There were multiple consolidations of municipalities in Finland during that period.

The Basic Education Act reform launched a new model for the system of special education. Before the reform, special education was divided into two tiers: part-time and full-time special education which were provided based on the needs of the pupil (Kirjavainen et al., 2016). The reform introduced a three-tier system: general support, intensified support, and special support. Every pupil is entitled to general support (Tier 1), which can mean part-time special or remedial education. Intensified support (Tier 2) is arranged to pupils who need regular support or different types of support simultaneously. The methods of intensified support are similar to the general support and the decision of transferring a pupil to intensified support is based on a teacher's pedagogical assessment and made by a multi-professional team. Special support (Tier 3) is arranged to pupils for whom intensified support is not sufficient.

There are several categories of reasons for special support. The categories include *emotional problem or social maladjustment*, *learning problems due to a specific language impairment, mild developmental delay, visual and hearing impairment,* and *other reasons.* Every pupil in special support gets an individual education plan which determines the means to offer support for the pupil. The plan also includes a decision about the level of individualization of the pupil. A special education pupil may follow the regular curriculum if it is suitable. However, she can have a partially individualized curriculum (PIC), which means that one or few subjects are individualized for the pupil. It is possible to have a mainly or fully individualized curriculum (MOFIC) which means that most subjects are individualized for the child. The decisions for MOFIC are often made for children with severe disabilities and at the age of 5 or 6. If MOFIC is not sufficient for the child, he can follow a modified curriculum. It means that the education is not separated into subjects as for most pupils. (Kirjavainen et al., 2016; Lahtinen and Lankinen, 2015, p. 230–236).

In Figure 2.1, we can see the developments in the shares of special education students over the past two decades. Before the latest reforms, the share of students in special support increased from 1995 to 2008. Most of the increase in special support occurred in milder categories such as *other reasons*

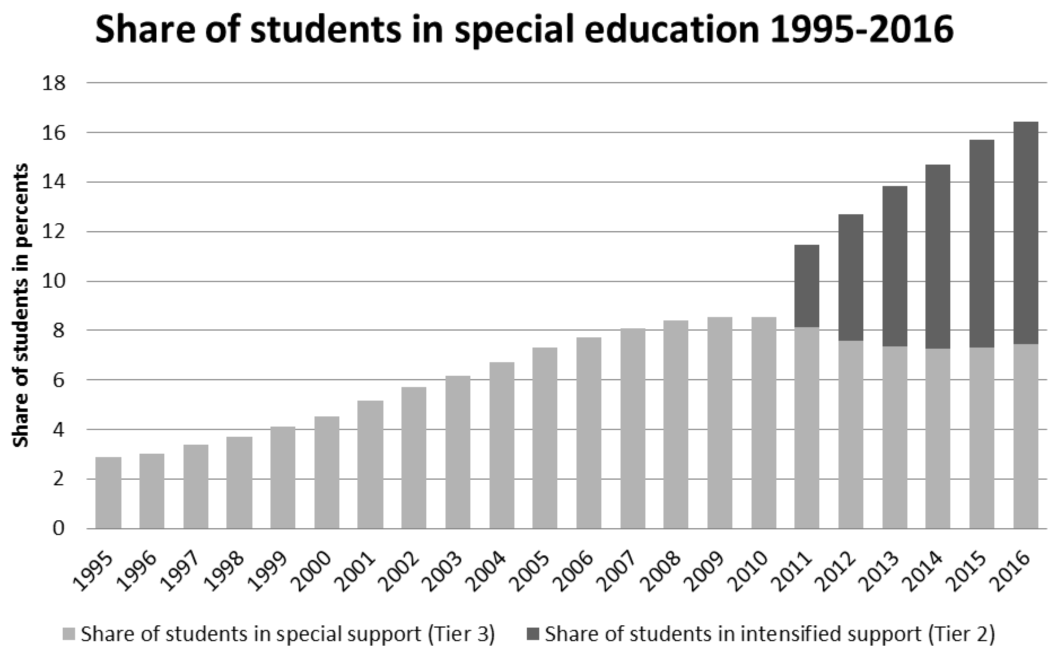

Figure 2.1: *Shares of students in special support (Tier 3) and intensified support (Tier 2). Author's rendering of data from Official Statistics of Finland (2016).*

(National Audit Office of Finland, 2013). After the reforms, the share of special support (Tier 3) students decreased slightly and has stayed constant since then. However, the share of students in the new tier, intensified support, has increased steadily in the 2010s. The increase in the share of special education students is a general trend in many developed countries and there is not one, fully explanatory, reason for this development (Graham and Jahnukainen, 2011). However, one possible reason for the increase in Finland is the fact that full-time special education became inclusive in a regular classroom and, thus, it became easier to assign students to full-time special education (Kirjavainen et al., 2016).

To sum up, special education in Finland has been in almost constant moulding since early 2000s with reforms and new curricula. Therefore, it is relevant to study whether the relative age effect has varied over time. A change in the relevant age effect can be interpreted as a change in the assignment criterion.

2.2. Special Education Research in Finland

Research on special education in Finland is mostly either qualitative or descriptive statistics. In the field of education, there are multiple studies that shed light on the state of special education in Finland and the effects of the recent reforms. However, there are not many studies that would find causal effects of special education or assignment to it. In this section, I review how special education in Finland has evolved in the 2000s and how it correlates with educational outcomes of students. I also present results on the geographical distribution of special education, how Finnish compulsory school principals perceive special education reforms, and how extra-funding for disadvantaged schools can improve educational outcomes.

National Audit Office of Finland (2013) studies how schools provided special education in Finland between 2001 and 2010. It focuses on trends in both part-time and full-time special education. Other questions of the report are: (1) how equally special education was provided, (2) does special education improve probability to continue for further studies, and (3) have the resources for special education increased over the period to match the increasing needs. Since the study covers the first decade of the 21st century, schools used the 2-tier system of special education during the research period.

The number of special education students between 2001 and 2010 increased by 60% from over 30,000 students to over 45,000 (National Audit Office of Finland, 2013). Similarly, the share of special education students increased from 5% to over 8% of all students. National Audit Office of Finland (2013) notes that the increase of special education students has mostly occurred in a dump category *other reasons* but also in *emotional problem or social maladjustment* and *learning problems due to a specific language impairment*. These are milder difficulties related to learning and behavior issues which are more subjective to diagnose (Kirjavainen et al., 2016). The increase of milder difficulties reflects on the curricula that students followed. The share of special education students who followed regular or PIC increased over the period, both from around 30% to around 40%. During the same period, the share of students with MOFIC decreased from over 40% to 22%.

National Audit Office of Finland (2013) finds that special education was not equally provided across the country and between different age groups. Younger generations have been provided more special education, which is consistent with the increase in the number of special education students. Furthermore, urban municipalities had higher shares of special education students. The fact that special education schools are usually located in cities explains only part of the difference. These regional differences raised concerns among the authors whether the regional equity that the Finnish Basic Education Act (628/1998) requires is fulfilled. Kirjavainen et al. (2014) also study regional differences in trends in special education between 2001 and 2010. They find that larger funds of the municipality is associated with larger number of special education students.

Special education students are less likely to continue for upper secondary education than students with no special support(National Audit Office of Finland, 2013). Kirjavainen et al. (2016) find that there are differences in the participation in upper secondary education between students with different levels of individualizations. While 94% of compulsory school leavers with a regular curriculum continue directly to further education, the shares for students with PIC or MOFIC are around 80% and 65%, respectively. Furthermore, almost 55% of the students with regular curriculum continue to general upper secondary school (academic track) and 40% to vocational school. This is completely different for students with individualization: only few percent of students with PIC go to the general upper secondary school. To conclude, the level of individualization correlates strongly with later educational outcomes.

Pulkkinen and Jahnukainen (2016) study the views of compulsory school principals and municipal education administrators on the funding reform of 2009, and the reform in the Basic Education Act of 2010. The researchers send questionnaires to 600 principals responsible for basic education and they receive replies from 335 participants. They conduct semi-structured interviews with seven municipal education administrators. According to the principals and administrators, special education decisions are made as the Act instructs with an emphasis on the pedagogical view. The focus on early intervention comes up in the surveys and interviews. The study indicates that the resources in part-time special education are particularly insufficient. However, administrators and principals do not consider the funding reform nor the special education reform significantly affecting the resources in special education.

Silliman (2017) studies how the "Positive Discrimination" policy affected the educational outcomes of low performing students in Helsinki, Finland. The funding policy was targeted to provide extra resources for disadvantaged schools to hire additional staff. Silliman finds that, in targeted schools, low-performing native students are 3 percentage points and students with immigrant backgrounds 6 percentage points less likely to drop out after middle school. Even though the resources were not particularly targeted for special education, they were meant to improve the educational outcomes of disadvantaged and low-performing students.

There are a lot of descriptive studies on special education and individualization in Finland. However, there exists a lack of research on the determinants of individualization decisions for special education students. This study is a starting-point for this research, studying the effect of relative school starting age as a one explaining factor. Prior Finnish research offers also interesting research questions. Since there are studies on recent reforms in special education it is relevant to study how these reforms have changed the relative age effect.

2.3. Special Education Research Abroad

The effects of special education programs are relatively mixed. Some studies find that special education improves student performance[8] whereas others do not find a significant effect (see Keslair et al.,

[8] See Cohen (2007); De Haan (2017); Hanushek et al. (2002); Henry et al. (2010).

2012) or a small effect relative to the costs (see Lavy and Schlosser, 2005). This is understandable since it is not expected that different programs in different countries would have similar effects. The identification strategies vary significantly among different studies, partly because the data available differ. Since there are different data sets and identification strategies, making definite conclusions on the issue is not easy. Some studies focus on the trends in numbers of special education students and the incentives affecting these trends (Iversen, 2013; Kwak, 2010).

One main difficulty to overcome in special education research is the issue of selection bias. Special education students are not comparable to regular students. Thus, to study the effects of special education, we need a quasi-experimental design. With a good design, the researcher can have two comparable groups: one which has received special education and one that has not. An example of this type of design is a study done by Keslair et al. (2012) who focus on the "Every child matters" program in the UK. They deploy their discovery that children are more likely to be assigned to special education if their peers at school perform better on average. Using this fact, the authors are able to compare groups of pupils who had scored in the same range in national exams but had different shares of special education children. They find that the program does not have a significant impact on children who are on the margin to be included in the program. They do not find any spillover effects for students who are not in the program. The researchers suggest few reasons for the zero-impact of special education. For example, children in special education may obtain a counterproductive stigma that might arise from lower expectation and shame.

In the US, Hanushek et al. (2002) show that special education raises achievement for students with disabilities. The authors follow students who move in and out of the special education program and compare their performance to their own expected trend. The researchers are able to find a significant improvement of individual performance in mathematics for the years that students are in a special education program. Similarly, Cohen (2007) finds that special education placement in Chicago Public Schools reduces the probability of absenteeism and dropping out of high school.

However, there is no clear effect on GPA. Setren (2016) finds that special education students benefit from charter schools in the Boston area. This is the case despite the fact that charter schools have a higher rate of integration of special education students to general teaching.

De Haan (2017) studies the effect of additional funds for low-ability pupils in the Netherlands. She uses a non-parametric bounds analysis to review the effects. With the non-parametric bounds analysis, De Haan is able to estimate a lower and upper bounds of the effects. However, with the design of the study there is no causal point estimate for the effect. The conclusion of the research is that adding funds for low-ability pupils in high school level can increase the schooling outcomes.

Iversen (2013) studies how an implementation of an educational reform in Norway affected the share of special education students. The reform, among other things, encouraged to enhance accountability in the governing of the municipality. The researcher finds that municipalities with a higher degree of implementation had significantly smaller increases in special education placements. Similarly, Kwak (2010) studies the effects of a funding reform in special education in California. Between 1996 and 1998, California reformed the funding from being based on the number of special education students to capitation which was based on total enrollment. Kwak finds that the increasing trend in the number of special education students in California stopped after the reform, whereas in the non-reform comparison group the trend kept increasing.

The international literature on special education is broad. However, since special education systems differ between countries it is important to study them separately. Thus, there should be more quantitative and causal research on the Finnish special education also.

2.4. Relative Age Effect

If two children were born close to each other but on different sides of New Year, they are close to each other in absolute age. Because of the school starting age rule, the child who was born in

December starts school a year earlier than the one born in January. Thus, the children have almost one year difference in age at the moment of their school start. This difference can impact children in many ways and it is called *the relative age effect*.

The school starting age of a child can affect the educational outcomes through multiple mechanisms. First of all, older pupils may have better skills at the moment of school start since they have had almost a year more time to acquire those skills. For example, Heckman (2006) emphasizes that "*early mastery of a range of cognitive, social, and emotional competencies makes learning at later ages more efficient*" and, therefore, this gap at the start may have long-lasting effects. At each point of evaluation, older pupils have accumulated their knowledge for a longer period and this gives them an advantage. In case of special education, this difference in ages at a moment of evaluation may induce that younger pupils are more easily assigned to special support. "*A common explanation is the inherent difficulty in distinguishing between maturity and ability when children are young and beginning formal schooling*" (Dhuey and Lipscomb, 2010).

Relative age affects the educational outcomes also through the peer effects. However, unlike the previous effect, it can impact children in both positive and negative ways. For example, being physically ahead of their peers, older students can feel more confident which in turn may improve their outcomes. For younger students, it may lead them to work harder to catch up with older peers or it may reduce their confidence and lead to poorer performances. A third mechanism of relative age effect is the concept of optimal school starting age. It is considered that a child can learn certain things at certain age more efficiently. The optimal school starting age undoubtedly depends on the individual, but there could be an age which would be, on average, best for children.

The above mentioned effects can interact with each other and they are difficult to separate. For example, since older students have accumulated more knowledge by the time they start school, they do better on average from the very beginning of school. This can lead to the aforementioned increase in confidence through the peer effects. Thus, the absolute age difference at every point is also a factor on peer effects. This difficulty to separate these effects comes up in the research design. For example, Black et al. (2011) separate the effect of school starting age from the age-at-test effect by analyzing the results of an IQ test which Norwegian men take when they turn 18. However, since the IQ test has a different cutoff date than the school starting does, it means that some individuals have had more schooling than others once taking the IQ test. The researchers are able to restrict the sample to the individuals who have already finished schooling, but one could argue that these effects can never be perfectly separated.

The existing literature on the relative age effect on special education is limited but there is a study from the United States which suggests that relatively younger children are more likely to get a special education status (Dhuey and Lipscomb, 2010). The researchers use a linear instrumental variable estimation to study relative age effect on receiving special education. They use a representative data which includes children from kindergarten until 10th grade. The researchers find that relative age effect is strongest for the category of learning problems. In contrast, there is no significant effect for hearing problems or orthopedic problems. This is consistent with an idea that relative age effect would exist in more subjective categories (Dhuey and Lipscomb, 2010).

There are also studies which find that relatively older students generally fare better at school. Kaila (2017) uses a regression discontinuity method to determine how relative age affects the GPA of middle school graduates and if relatively older people do better in the general upper secondary school. He finds that students born right after New Year have on average a 0.15 grade points higher GPA and are more likely to be admitted to and graduate from general upper secondary school compared to the students almost a year younger. The evidence from Sweden suggests similar results (Fredriksson and Öckert, 2014).

Relatively younger children are more likely to be diagnosed with ADHD too (Elder, 2010; Sayal et al., 2017; Schwandt and Wuppermann, 2016). Elder (2010) finds that children's birth date influenced the teachers' assessment much stronger than the parents' assessment. Elder writes, "*Discontinuities* *around eligibility cutoffs in teacher reports of hyperactivity and inattentiveness are four times larger than the corresponding discontinuities based on parent reports. These patterns suggest that teachers' opinions of children are the key mechanisms driving the relationship between school starting age and ADHD diagnoses.*" The researcher concludes that teachers most likely compare children to their classmates who can have substantially large age differences at early grades. This is one possible reason for the relative age effect for the assignment to special education. If teachers compare children to their peers who are within the same grade rather than same age, younger children may seem less developed than they actually are.

Even if relative age effect is a well-studied phenomena, there is not much research on its effect on special education. It is important to study this effect in different contexts to be able to generalize the effect and form better theories of it. This thesis takes this study to a new country. Furthermore, I use a register data which increases the reliability of my results.

3. Methodology

3.1. Research Design

An ideal way to study the effects of relative age for special education, as in many other cases, would be an experiment with randomization. If it was ethically plausible we could run an experiment in which we randomize a group of same-aged children (say, born in January) so that half of them would start school a year earlier than the other. From this experiment, we could estimate the average treatment effect of being a year older by comparing the shares of children in special education between the two groups. However, this kind of experiment could be considered unethical and the experimenter would have to wait for the results of long term outcomes.

As Angrist and Pischke (2009, p. 4–5) point out, it can be fruitful to consider the ideal experiments even if running them might not be an option. It helps to find the relevant research questions or consider suitable methods for the research. In our case, it helps to find suitable groups which we can compare to each other. Children born around New Year can be compared to each other to study the effect of relative age.

To formalize our research design, we use the concept of potential outcomes. In this formalization, I follow the notation of both Angrist and Pischke (2009, p. 13–15) and Imbens and Lemieux (2008). We can think of two potential outcomes for each individual. In our case, this can mean (1) having completed a regular curriculum by the end of basic education or (2) having not completed it. These two alternative outcomes for an individual i, I denote as $Y_i(0)$ and $Y_i(1)$. Since I study the relative age effect, the treatment in my study is whether a child is old compared to her peers or not. Thus, I denote $OLD_i = 0$ if an individual i did not receive treatment and $OLD_i = 1$, respectively, if she did receive the treatment.[9] We can formulate the observed outcome in the following manner:

$$Y_i = (1 - OLD_i) \cdot Y_i(0) + OLD_i \cdot Y_i(1) = \begin{cases} Y_i(0) & \text{if } OLD_i = 0, \\ \\ Y_i(1) & \text{if } OLD_i = 1. \end{cases} \tag{1}$$

Unfortunately for research, it is impossible to observe both outcomes for one individual. Therefore, we have to concentrate on the differences between averages of two subpopulations. This average treatment effect is the key to causality. In principle, if the two groups are as good as randomly assigned, the difference between the averages is the causal effect of the treatment.

Since we want to study the causality of relative age on special education, a simple comparison of averages or ordinary least squares regression is not a suitable option. If we compared the frequencies of students with individualized curriculum between the old and young in relative age, we would not see the causal effect of relative age on the likelihood of having an individualized curriculum. There would be a selection bias on the relative age of the students so they would not be randomly assigned. To illustrate this, I follow the notation of Angrist and Pischke (2009, p. 14):

$$\begin{aligned} E[Y_i|OLD_i = 1] - E[Y_i|OLD_i = 0] = {} & E[Y_i(1)|OLD_i = 1] - E[Y_i(0)|OLD_i = 1] \\ & + E[Y_i(0)|OLD_i = 1] - E[Y_i(0)|OLD_i = 0]. \end{aligned}$$

The left-hand side of the equation is the difference in the outcome that we can observe in the data. The first term of the right-hand side is the causal effect on the relatively old. The second term is the selection bias, that is the characteristic difference between the treatment and control group, the young and the old.

It is unlikely that the old and the young would be characteristically similar groups, meaning that

[9] I define the exact meaning of being old (ie. $OLD_i = 1$) in subsection 3.4 as I now concentrate on more universal principles of the research design.

there would be no selection bias. If a child starts school later than assigned to, they usually have a reason for doing that. Thus, they may fare worse at school and be more likely to receive special education. Similarly, students who are held back a year are more likely to have an individualized curriculum. The selection bias applies in the other end too; children who start school earlier than assigned to are tested if they are mature enough. Thus, they might be expected to do better than average at school. Due to these selection biases, a naive comparison might estimate an opposite effect than we hypothesize: relatively older students would seem to be more likely to have an individualized curriculum. Therefore, we need a more sophisticated design and a quasi-experimental method.

In the case of this study, we can exploit the discontinuity of school starting age. As mentioned in subsection 2.1., children in Finland start school in August of the year they turn seven. This means that children who are born right before New Year are (in most cases) almost a year younger than children who are born after New Year when they start school.

A method that utilizes a discontinuity is called regression discontinuity (RD) design (see Angrist and Pischke, 2009, p. 251–267). Under certain assumptions, we can analyze RD design like a randomized experiment (Lee and Lemieux, 2010). Since the cutoff date for starting school is not strictly forcing, I use a design called fuzzy regression discontinuity (FRD). This method takes into account that some children start the primary school a year earlier than their birth date assigns and others a year later. Next, I present the theoretical framework of FRD. Finally, I continue to a concept of local average treatment effect (LATE) which closely relates to FRD.

3.2. Fuzzy Regression Discontinuity

The aim of regression discontinuity (RD) is to utilize an artificial jump on the treatment variable OLD_i to estimate its effect on the outcome variable Y_i. In RD setting, there is an assignment variable A_i[10] which has a cutoff point c in which the jump in the probability of the treatment OLD_i occurs. If the cutoff is binding, the probability jumps from 0 to 1. This type of RD is called sharp regression discontinuity. When the jump at the cutoff is smaller, the setting is called fuzzy regression discontinuity.

Fuzzy regression discontinuity is basically an instrumental variable (IV) design. As an IV setup, it also contains two stages. The first stage is to estimate the probability of complying the cutoff rule. In the second stage, we use this probability to estimate the actual effect of the discontinuity.

The first essential property of the setup is that the probability does have a discontinuous jump. Without this jump, there is no way to use it as an instrumental variable. Mathematically this is written in the following way:

$$\lim_{a \downarrow c} \mathrm{P}(OLD_i = 1 | A_i = a) \neq \lim_{a \uparrow c} \mathrm{P}(OLD_i = 1 | A_i = a). \tag{2}$$

This jump is illustrated in Figure 3.1, which is an example of a first stage of FRD. The figure shows how the probability of receiving treatment depends on the assignment variable a. As a increases, the probability of receiving treatment increases too, in this case linearly. There is a jump in the probability at $a = 2$. Thus, the cutoff point c for the assignment variable is 2. In the first stage of FRD, the jump of the treatment effect is estimated: (Imbens and Lemieux, 2008)

$$\theta_1 = \lim_{a \downarrow c} E[OLD | A = a] - \lim_{a \uparrow c} E[OLD | A = a]. \tag{3}$$

The second stage of FRD is to estimate the effect of the treatment around the cutoff point c. This

[10] There are multiple terms for this variable in the literature. In this thesis, I call it an assignment variable or an instrument variable depending on the context we look at it.

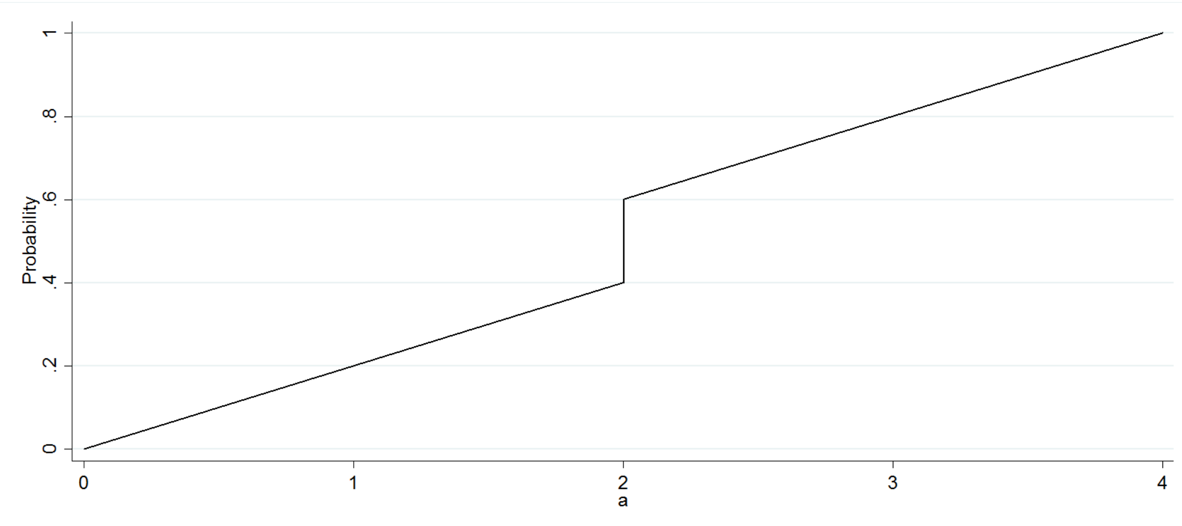

Figure 3.1: *An illustration of discontinuity in the probability of assignment. In this case, the discontinuity is at point a=2.*

average effect of the treatment is formally the following:

$$\theta_2 = \lim_{a \downarrow c} E[Y|A = a] - \lim_{a \uparrow c} E[Y|A = a]. \tag{4}$$

The actual causal effect can be interpreted as a Wald estimand. It is the average treatment effect divided by the average jump in the probability of the treatment: $\tau_{FRD} = \theta_2/\theta_1$ (Imbens and Lemieux, 2008). In Figure 3.2, the solid line represents a curve that could be seen from the data. It is called the actual outcome since it is what we would actually see from data. The two dashed lines illustrate the curves of potential outcomes for both the treatment and the comparison groups. The height of the jump in the solid line is not the causal effect of the treatment but represents θ_2. The causal effect is the distance between the two dashed lines at point $c = 2$. The actual causal effect is larger than the jump in the solid line since the treatment status does not change for everyone at the cutoff. Again, if the RD was sharp, the jump would equal to the causal effect but that is not the case for FRD. Thus, if we did not use the IV setup, the actual effect would be underestimated.

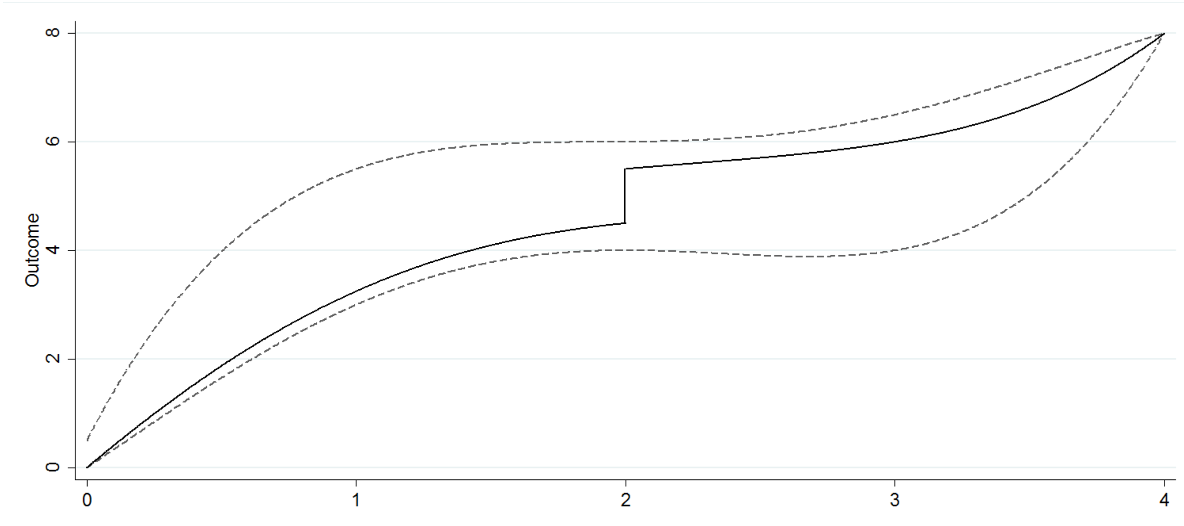

Figure 3.2: *An illustration of potential outcomes for the comparison and treatment groups in dashed lines. The actual outcome with a discontinuity in a solid line.*

In my research design, I use the birth date of an individual as the instrument for FRD. I simplify the instrument into a binary variable so that $A_i = 0$ if an individual is born before New Year (the one nearest to the birth date) and $A_i = 1$ if an individual is born after New Year. I have also a binary variable for the treatment: $OLD_i = 0$ if an individual is among the youngest of the class and $OLD_i = 1$ if an individual is among the oldest. I have more exact definitions of these variables in subsection 3.4 which includes a detailed description of the estimation strategy.

There are certain considerations regarding FRD. Since the estimand obtained with FRD can be interpreted as a local average treatment effect (LATE), the assumptions of LATE must be satisfied. I present the LATE assumptions in the next subsection. Another issue that has to be considered in RD designs is the selection of bandwidth. Technically, we are estimating the effect of a year in relative age, meaning the gap between being born right before and right after New Year. However, there is a question of sufficient size of the sample that is included in the regression. Thus, there is a trade-off between the distance from the actual cutoff and the sample size. In theory, the bandwidth does not have to be same on both sides of the cutoff but I use a symmetric bandwidth. It is more relevant to study whether the effect varies with different bandwidth choices, than to choose one single bandwidth.

There are tests that can be used for choosing an optimal bandwidth. For example, we could use cross-validation or a standard F-test to compare the fit of a regression with different bin widths and bandwidths (Lee and Lemieux, 2010). When using these methods, the optimal bandwidth depends on the chosen outcome variable. This would lead to different samples for the different outcome variables. Thus, I choose to use a constant bandwidth of 30 days on each side of the cutoff and I do a robustness analysis on the bandwidth choice. Since the birth date data is discrete, it contain natural bin widths (the date). Thus, it is reasonable to focus only on the robustness of the results with respect to bandwidth.

3.3. Local Average Treatment Effects

Since our fuzzy regression discontinuity design has a binary instrumental variable (A_i) and a binary treatment variable (OLD_i), we can interpret the Wald estimand as a local average treatment effect (Lee and Lemieux, 2010). The design has to satisfy the assumptions of the LATE theorem to be able to treat our Wald estimand as a causal effect. Angrist and Pischke (2009, p. 155) put LATE theorem in words as follows:

> This theorem says that an instrument that is as good as randomly assigned, affects the outcome through a single known channel, has a first stage, and affects the causal channel of interest only in one direction can be used to estimate the average causal effect on the affected group.

We can separate the theorem into four conditions for the instrument, to be able to use it to estimate the average causal effect. In this subsection, I present these four assumptions. Afterwards, in subsection 5.1, I argue that the assumptions are satisfied in the setting of this study.

The first assumption is called *Independence* which is mathematically formulated in the following way: $\{Y_i(0), Y_i(1), OLD_i\} \perp\!\!\!\perp A_i$, for $OLD_i = \{0, 1\}$ (Angrist and Pischke, 2009, p. 152). In other words, the instrument is independent of the outcome and the assignment variable.

The second assumption of the LATE theorem is called *Exclusion*. Its specific way of writing is $Y_i(OLD_i, A_i = 0) = Y_i(OLD_i, A_i = 1)$ for $old = 0, 1$(Angrist and Pischke, 2009, p. 153). Exclusion restriction says that the only way the instrument affects the outcome is through the treatment variable. Again in our case, this means that the absolute age of a child affects the probability only through the relative age effect.

The existence of a *first stage* is the third assumption of the LATE theorem. The precise formulation of it is $E[OLD_{1i} - OLD_{0i}] \neq 0$ (Angrist and Pischke, 2009, p. 155). In the case of our FRD, this relates back to Figure 3.1 and the jump in the probability of the assignment. This enables us to use the jump as a first stage estimator.

The last assumption for the theorem is *Monotonicity*. It is stated in the following way: $OLD_{1i} - OLD_{0i} \geq 0 \quad \forall i$, or vice versa (Angrist and Pischke, 2009, p. 155). This assumption means that if the treatment affects an individual, it affects all the individuals in the same way. In our study, this means that even if not everyone decides to start school a year later if being born after New Year, being born after New Year can only affect the decision to one direction: going to school later. To explain this, I separate four types of individuals: *compliers*, *always-takers*, *never-takers* and *defiers* (Angrist and Pischke, 2009, p. 158–160). Compliers are the ones who follow the rule of the instrument for the treatment by going to school the year they are assigned to. If they are born before New Year, they go to school early and, respectively, if they are born after New Year, they start school a year later. Never-takers do not take the treatment regardless of the instrument. They start school early regardless of the birth date. Similarly, always-takers "take the pill" regardless of the instrument. In our study, always-takers start school later even if they were born before New Year. Defiers are individuals who work against the logic of the instrument. If the instrument assigns them to treatment, they do not take it but if they are not assigned to the instrument, they will take it. The Monotonicity assumption does not allow defiers. In our study, defiers would start school later if they were born in late December and earlier if they were born in January.

3.4. Estimation

In this subsection, I explain the estimation strategy in detail. I present the equations which I use for the estimation and explain why they are chosen. I use FRD design for the estimations but I run them also in reduced form. The reduced form illustrates why I choose to use FRD design. Before moving to these equations, I define the essential variables for the estimation.

First of all, I define a birth date variable, BD_i, which is normalized around the cutoff point, New Year. To obtain this variable, I subtract the nearest January 1st from individual's birth date. BD_i tells the distance from the closest New Year and on which half of the year an individual was born. BD_i gets negative values if a child was born between July and December and positive values if she was born between January and June. For the estimation, I pool all individuals in the sample by their birth date, so that I am able to estimate the effect of the cutoff date.

The assignment variable, $After_i$, is a dummy variable which tells whether an individual i was born before or after the cutoff. I define it the following way:

$$After_i = \begin{cases} 0, & \text{if } BD_i < 0 \\ 1, & \text{if } BD_i > 0. \end{cases} \tag{5}$$

For defining the relative age, I use this normalized birth date BD_i but I also take into account that some students start school earlier or later than usual. Thus, I add a variable GA_i, which is the age of an individual i at the end of the year of middle school graduation. For most students, $GA_i = 16$, since they turn 16 during the year they graduate from middle school. Now, I can define relative age the following way:

$$OLD_i = \begin{cases} 0, & \text{if } (BD_i < 0 \wedge GA_i = 16) \vee GA_i < 16 \\ 1, & \text{if } (BD_i > 0 \wedge GA_i = 16) \vee GA_i > 16. \end{cases} \tag{6}$$

In words, relative age, OLD_i, equals zero if student i was born between July and December and turns 16 the year she graduates from secondary school or is younger than that. Respectively, OLD_i equals one if she was born between January and June and turns 16 the year she graduates or is older than that.

I define the reduced form regression in the following way:

$$Y_i = \alpha_1 + \alpha_2 After_i + f_1^k(BD_i) + g_1^k(BD_i After_i) + \boldsymbol{\alpha X_i} + \delta_i. \tag{7}$$

In this equation, Y_i is the chosen outcome variable for an individual i, and $BD_i After_i$ is an interaction term of the birth date and the assignment variable, $\boldsymbol{X_i}$ is a vector of control variables, and δ_i is the standard error term. The control variables are include parents' education and income, whether a student lives with one or both parents, gender and year fixed effects. The interaction term is needed to separately estimate the polynomial on both sides of the cutoff. If there was no interaction term, the plotted lines would be parallel on both sides of the cutoff. Functions f_1^k and g_1^k are kth order polynomial functions of assignment variable and the interaction term. In my main estimation, I use only linear functions since RD effects with a high order of polynomials are usually not as accurate as linear or quadratic forms of the assignment variable (Gelman and Imbens, 2017).

I define the two stage equations of FRD design in the following way:

$$OLD_i = \beta_1 + \beta_2 After_i + f_2^k(BD_i) + g_2^k(BD_i After_i) + \boldsymbol{\beta X_i} + \varepsilon_i \tag{8}$$

$$Y_i = \gamma_1 + \gamma_2 OLD_i + f_3^k(BD_i) + g_3^k(BD_i OLD_i) + \boldsymbol{\gamma X_i} + \omega_i. \tag{9}$$

The first stage estimates the probability of being old in the classroom using the assignment variable as an instrument. In the second stage, I use this relative age variable to determine its effect on a chosen outcome variable. Functions f_2^k,f_3^k, g_2^k, and g_3^k are polynomial functions corresponding to f_1^k and g_1^k.

The reduced form is supposed to give a smaller effect since it does not take into account the fuzziness of the discontinuity. It only estimates the jump seen in the data, as explained in subsection 3.2. whereas the two stage method divides this estimate with the probability from the first stage to obtain τ_{FRD}.

For the estimation, I use the triangular kernel function to put greater weight on individuals whose birth date is closer to the cutoff point. The equation to define the weights is the following:

$$weight_i = \begin{cases} 1 - \frac{|BD_i|}{h}, & \text{if } |BD_i| \leq h \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

In the equation, h is the chosen bandwidth on both sides of the cutoff. The choice of kernel function is not empirically essential since the differences between different weights do not usually affect the estimates (Fan et al., 1997). However, it has been shown that the triangular kernel is optimal for estimating local linear regressions at the boundary (Fan and Gijbels, 1992). For robustness, I show the estimation results using a uniform kernel with different bandwidths in subsection 5.3.

4. Data

The data I use is offered by the Statistics Finland. I have data from all compulsory school leavers between 1998 and 2014. This data set is called National joint application register. The data contain information on the school that the student leaves, students' mother tongue, GPA, individualization of the curriculum, and information on where they apply for secondary education. In addition to National joint application ragister, I use Finnish Longitudinal Employer-Employee Data (FLEED) to connect parents' educational and income information with the students. I have information from separate data sets on students' immigration status, birth date, and whether they live with their parents.

There is no explicit information on the special education status in the data. Thus, I have to use the individualization of the curriculum as a variable of interest. We can see from Table A.1 in Appendix, that 58% of 9th grade students in special support have an individualized curriculum on average. Special education students with an individualization have most likely more difficulties in their studies compared to those who follow a regular curriculum.

The variable of interest is called the *level of individualization* of a student. The variable tells whether the student finished a *regular curriculum*, *partially individualized curriculum* (PIC), *mainly or fully individualized curriculum* (MOFIC) or if she had a *modified curriculum.* The same variable tells if a student dropped out from compulsory school without a diploma. This means that we do not know the type of individualization of the dropouts. As mentioned previously, a special education student may have a regular or an individualized curriculum depending on their needs. My main interest is the PIC students since the decisions made for this category are more subjective and relative age could be expected to have the most effect on that category. Additionally, most of the increase in special education students has occurred due to milder learning difficulties, which comes up in the trend of the students with PIC. (Kirjavainen et al., 2016).

Table 4.1: *Middle School Graduates*

Year	Regular Curriculum	Partially Individualized Curriculum	Modified Curriculum	Mainly or Fully Individualized Curriculum	Drop out	Total
1998	60305	381	95	953	265	61999
1999	60607	349	100	1040	338	62434
2000	59120	608	92	958	421	61199
2001	56613	651	93	1029	348	58734
2002	54224	780	106	979	253	56342
2003	53155	986	119	1168	256	55684
2004	55680	1191	117	1214	285	58487
2005	55221	1457	53	1190	43	57964
2006	56981	1866	125	1434	48	60454
2007	56443	2147	56	1324	87	60057
2008	57569	2332	73	1337	103	61414
2009	58843	2645	57	1337	98	62980
2010	58247	2618	39	1289	129	62322
2011	56779	2597	41	1261	171	60849
2012	54551	2550	49	1278	177	58605
2013	54173	2351	49	1003	186	57762
2014	60641	2588	47	911	114	64301
Total	969152	28097	1311	19705	3322	1021587
	94.9%	2.8%	0.1%	1.9%	0.3%	

Notes: Number of middle school graduates in different categories for 1998–2014.

Table 4.1 shows how numbers of students in different categories have changed over time. As we can see, the total number of middle school graduates has been around 60,000 over the period. The lowest value was reached in 2003 with 55,684 middle school graduates and the highest in 2014 with 64,301 graduates. The number of students with PIC has increased over the period. From 1998 to 2014, the number of students with PIC has increased from 381 to 2588. This growth means more than six-fold increase and almost a 12% annual increase. In contrast, the number of students with MOFIC has stayed stable over time with around 1000 students. The number of drop outs and students with modified curriculum has, on average, decreased over time but there is a large yearly variation in both of them.

Figure 4.1 depicts the trends with different levels of individualization as shares of the total number of students. The upper graph shows how the share of students with regular curriculum has varied over time. In 1998, over 97% of middle school graduates had a regular curriculum. Their share has decreased over time with a considerable gap between 2005 and 2006. By 2012, the share of regular curricula had decreased to 93% from which it has bounced back to 94% by 2014. The lower graph depicts the shares of students with individualized curricula. The solid dots present the share of students with MOFIC and show that their share has stayed at around 2%. The circles show the shares of PIC and it tells a different story. Before 2000, the share was around 0.5% from which it started to increase and by 2005, the share of PICs had exceeded the share of MOFICs. In 2009, the share of PICs reached 4% since when it has stayed relatively constant.

Unfortunately, there is no information on the timing of the individualization decisions. It would be valuable to see if relative age effect varied at different school grades. On an aggregate level, we know that the number of students in special support increases through compulsory schooling (National Audit Office of Finland, 2013). However, there is a jump upwards in the trend from the 6th grade to 7th when students move from primary school to the lower secondary school. So, children are assigned to special support quite evenly through compulsory school and the shares of students in

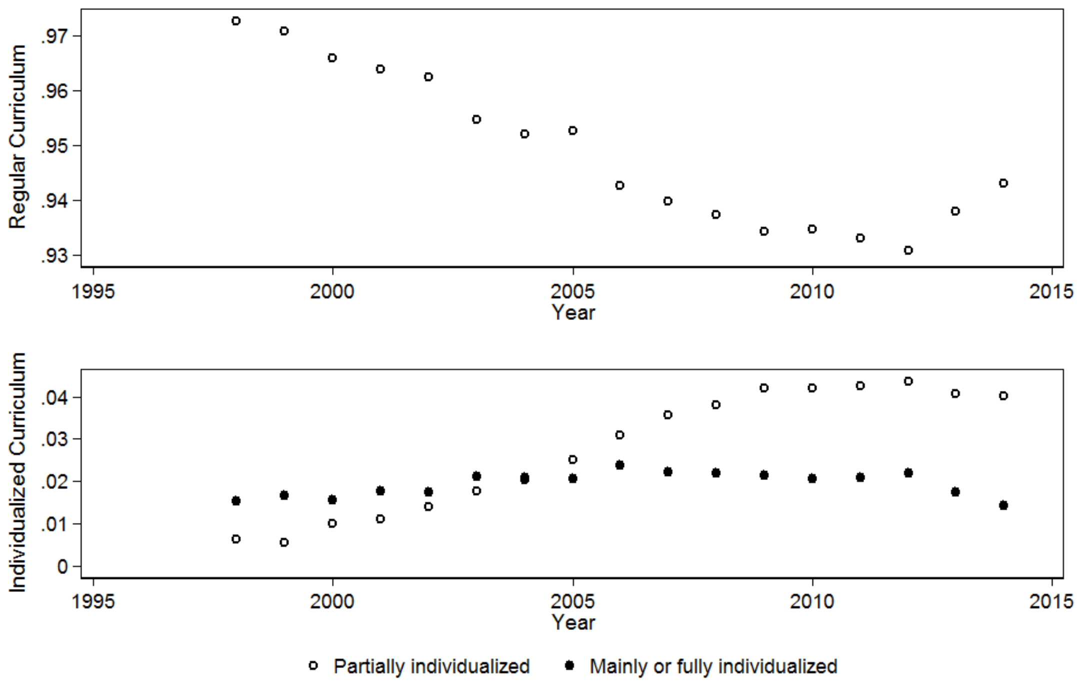

Figure 4.1: *Shares of students in different levels of individualization. The share of students with regular curriculum are plotted in the upper graph. Shares of students with PIC and MOFIC are in the lower graph. Author's rendering of data from Official Statistics of Finland.*

special support are highest for final grades.

Another restriction of the data is that there is no information of school starting year. Therefore, I cannot separate students who have started school a year later from ones who have repeated a class. This leads to overestimation of the number of students who have started school a year later and leads to an upward bias of the instrument estimation. However, I cannot determine if a child started school a year earlier than assigned to and was held back during primary school. These students lead to a downward bias and the total bias is the sum of these two biases.

To obtain the estimation sample, I restrict the sample of my data in the following way. First of all, my main data set, National joint application register, contains information on two groups: (1)

all applicants to secondary education and (2) students who finished compulsory school but did not apply for secondary education. My interest is solely in the students who have finished compulsory school that year. Thus, I drop all applicants who applied for secondary education but had finished compulsory school earlier. I drop all the students who are born abroad, whose parents are not Finnish, and are born on January 1, because they are over-represented for that birth date. It is typical that children whose birth date is unknown are set to have it on January 1.

Table 4.2 represents the summary statistics of the data by different level of individualization. Parents' education is classified in the following way: the values range from 3 to 8 in which 3 means the upper secondary education, 5 is short-cycle tertiary education, 6 is bachelor's or equivalent level, 7 is master's or equivalent level, and 8 is doctoral or equivalent level. In the regressions, I use the fixed effects of each of these levels. Earnings are the household's after-tax income including all transfers. The rest are dummy variables. As we can see from the data, individualization correlates with lower education, earnings, and probability to live with parents. Students with individualization are also predominantly males.

Table 4.2: *Summary Statistics*

	Regular Curriculum	PIC	MOFIC
Mother's Education	4.26	3.49	3.47
	(1.50)	(1.05)	(1.04)
Father's Education	4.28	3.45	3.45
	(1.58)	(1.04)	(1.06)
Mother's Earnings	25705	23094	19588
	(42121)	(21904)	(26469)
Father's Earnings	29010	23992	21136
	(56224)	(21279)	(33100)
Lives with Father	0.74	0.65	0.62
	(0.44)	(0.48)	(0.49)
Lives with Mother	0.93	0.88	0.86
	(0.26)	(0.33)	(0.35)
Lives with both Parents	0.69	0.58	0.54
	(0.46)	(0.49)	(0.50)
Female	0.50	0.38	0.34
	(0.50)	(0.49)	(0.47)

Notes: Summary statistics of the control variables by different outcome variable. The table shows the mean and the standard deviation in parentheses. Living with either parent and female are all dummy variables. Education variable ranges from 3 to 8 in which 3 means the upper secondary education, 5 is short-cycle tertiary education, 6 is bachelor's or equivalent level, 7 is master's or equivalent level, and 8 is doctoral or equivalent level. Earnings are the household's after-tax income in euros including all transfers.

5. Empirical Work

This section is split into three subsections. First, I examine the validity of the research design. I go through each of the LATE assumptions to justify my causal interpretation of the results. Then, in the second subsection, I present the main results both graphically and analytically. I present results considering the main questions of the thesis: is there a relative effect on the probability of having an individualized curriculum, how does it appear in different subgroups, and has the effect evolved over time. Lastly, in the final part of the section, I present further checks to convince that my results are robust and consistent.

5.1. Validity of the Estimation

In our case, the first LATE assumption, independence, means that a child's birth date is independent of the potential outcome, the level of individualization, and the potential treatment assignments. The latter means that a child's probability distribution to follow the school starting age rule is independent of the birth date. The former means that birth dates around New Year are as good as randomly assigned. However, there is evidence from certain cases that there has been manipulation of the birth date. For example, when Australia introduced a $3000 "Baby Bonus" for children born on or after July 1, 2004, there were more births on that July 1 than for any single date in the previous 30 years (Gans and Leigh, 2009).

Since manipulating birth date is possible at least to some extent, we need to discuss whether that threatens the validity of our study. One way to study this question is presented by McCrary (2008). In Table 5.1, I do a continuation test for the number of students on both sides of the cutoff date. I use equation 7 for the estimation and the table presents the value of α_2. The standard errors are large, so we do not find a statistically significant effect on the discontinuity of the number of students. Furthermore, once we proceed to graphical analysis, we see that there are some outliers

Table 5.1: *McCrary's Test*

	(1)	(2)
Discontinuity	129.9	0.0566
s.e.	(141.6)	(0.0584)
Constant	2524	7.825
s.e.	(116.7)	(0.0492)

Notes: A version of McCrary's test which analyzes the continuity of the sample used. In column (1), there are all the individuals born around New Year in a bin for each birth date. In column (2), is the logarithm of the sums. There is no evidence of discontinuity at New Year.

which can further confuse the analysis.

Dickert-Conlin and Elder (2010) study whether there are discontinuities in birth dates around the school starting cutoff dates in the United States. Their hypothesis was that mothers would optimize births to happen just before the cutoff date so that children could start kindergarten earlier. This would lead to savings since parents would not need to pay for daycare. This is not likely to be the case in Finland since daycare is significantly subsidized. However, it has been found that relatively older children do better in school and are more likely to apply to and graduate from upper secondary school (Kaila, 2017). Thus, parents could have the motive to delay the birth of a child past New Year. This is a debatable concern since there is evidence that even if the oldest of the classroom do better academically the gains may diminish in the long run and the prime-age earnings are unaffected (Fredriksson and Öckert, 2014). Fredriksson and Öckert (2014) find that higher school starting age, on average, causes lower discounted lifetime earnings. Older students enter the labor market later and are behind in accumulating work experience compared with the younger peers.

Another way to study the independence of birth dates is to take a look at their distribution around New Year. Figure 5.1 shows that birth dates are mostly quite evenly distributed. There are some clear outliers for the Christmas dates (Dec 24–26), the Finnish Independence day (Dec 6) and New

Year (Jan 1). Lower shares of students for those dates is because no elective premature births, like cesarean sections, are arranged for those dates. However, there does not seem to be any discontinuities in the number of students with different birth dates otherwise (see Figure A.1 in Appendix for a similar graph without the outliers).

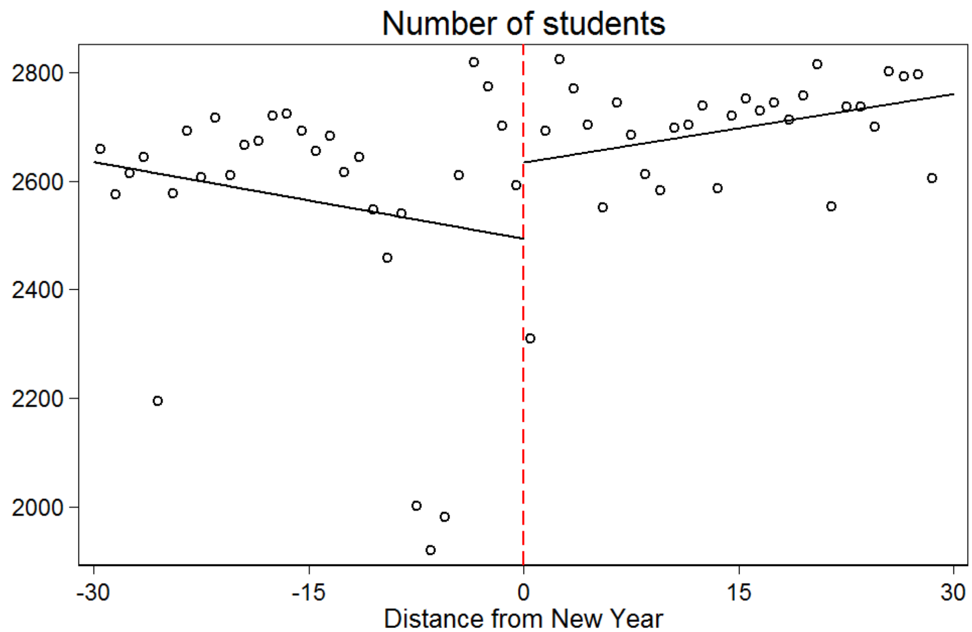

Figure 5.1: *Graphical view on the number of students born in around New Year. The lines are local linear estimations on both sides of the cutoff. This is the same model as presented in equation* (7) *with the number of students as an outcome variable. There are clear outliers for the Christmas dates (Dec 24–26), Independence day (Dec 6) and New Year (Jan 1). I show the same figure without the outliers in Appendix. Author's rendering of data from Official Statistics of Finland.*

The exclusion assumption means in our case that birth date only affects the outcome through the assignment variable. In practice, this means that birth date only impacts the probability of individualization by affecting the school starting age. There is evidence that the season of birth may affect cognitive and psychological development (Barak et al., 1995). However, the effect of season of birth is not a concern, because I study children born around New Year and they are all born in the same season. When I use the sample for the whole year, this can threat the exclusion

Table 5.2: *Control Variables*

	Mother's Education	Mother's Income	Father's Education	Father's Income
Discontinuity	-0.004	208.3	-0.015	-470.5
s.e.	(0.0182)	(422.4)	(0.0197)	(501.8)
Sample size	130612	159855	124295	159855
	Lives with Mother	Lives with Father	Lives with both Parents	Gender
Discontinuity	-0.018	-0.005	-0.006	0.007
s.e.	(0.0030)	(0.0049)	(0.0052)	(0.0055)
Sample size	159855	159855	159855	159855

Notes: Discontinuity tests for the control variables. The regression equation used is equation (7) with each control variable as an outcome. In addition to the coefficient, there is the standard error for each variable in parentheses.

assumption. Another problem that threatens the exclusion assumption is that New Year is also a cutoff for other instances than school starting. For example, in many sports and other hobbies children are separated into groups based on their birth cohort. Hobbies can have an impact on children's school outcomes and the possibility of this violation of the assumption cannot really be excluded.

Another way to study the exclusion restriction is to take a look at the control variables. If the instrument affects the outcome only through the treatment variable, it means that all the control variables should be continuous around the discontinuity point. In Table 5.2, I have a discontinuity test for each control variable. As we can see, there is evidence for a discontinuity in the probability of living with the mother but other control variables seem to be continuous. To decrease the possible problem that might arise from the discontinuity, I add the variable of living with the mother as a control in the regressions I run. In Figure 5.2, I present the average values of the control variables for children born in around New Year. The regression lines seem relatively smooth, and there are no visually remarkable jumps. To study the smoothness of the control variables further, I run continuation tests on them. They are control variables for the regressions, so the correlation between the relative age and the outcome variable does not come through them.

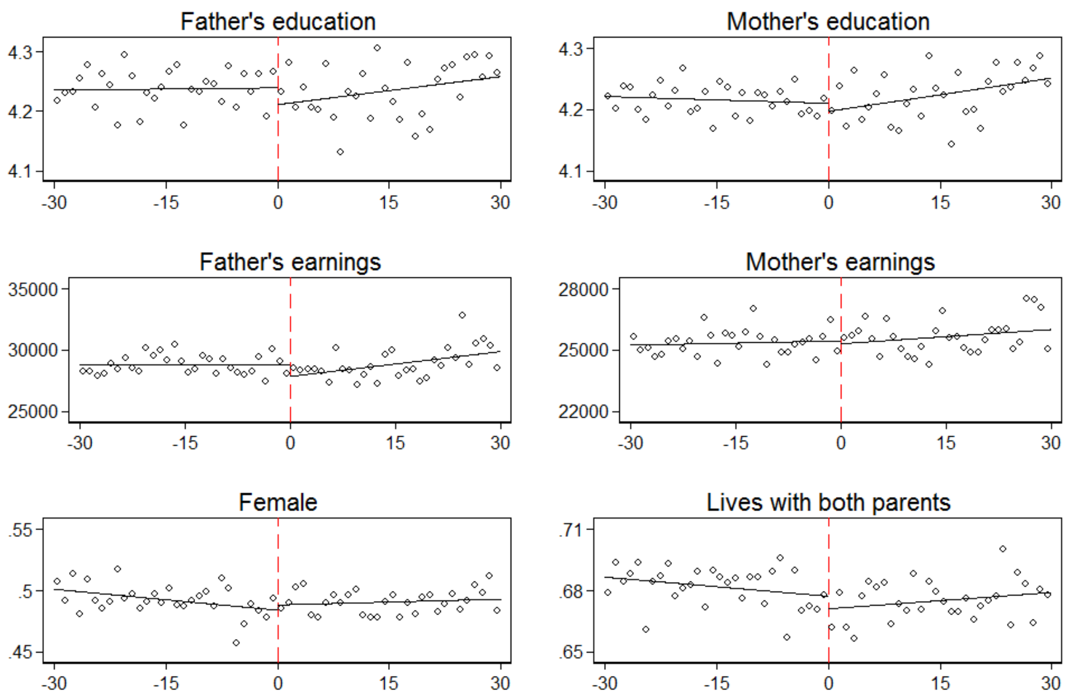

Figure 5.2: *Graphical view on the average values of the control variables for students born around New Year. I present local linear estimations on both sides of the cutoff. This is the same model as presented in equation (7) with a given control variable as an outcome. Author's rendering of data from Official Statistics of Finland.*

The third assumption is about the existence of a first stage. In our study, it means whether there is a jump at New Year in individual's probability to be relatively old at the school class. Figure 5.3 illustrates this jump of the first stage in our study. As we can see, there clearly is a jump at the cutoff suggesting that the first stage exists. I show the coefficients for the first stages in the next subsection.

What comes to the final assumption of monotonicity, it is more challenging to confirm that this assumption is fully met. Since we only see one outcome for an individual (either comply the instrument or not), it is not entirely possible to determine to which category an individual belongs to. For example, an individual who is born in January and is among the oldest of the class could be either a complier or an always-taker. However, it is not really reasonable to expect that an individual would start school a year late if they were born in December but in the contrast would start it a year early if they were born in January. Thus, defying is not really plausible behavior and we can trust that the monotonicity assumption holds.

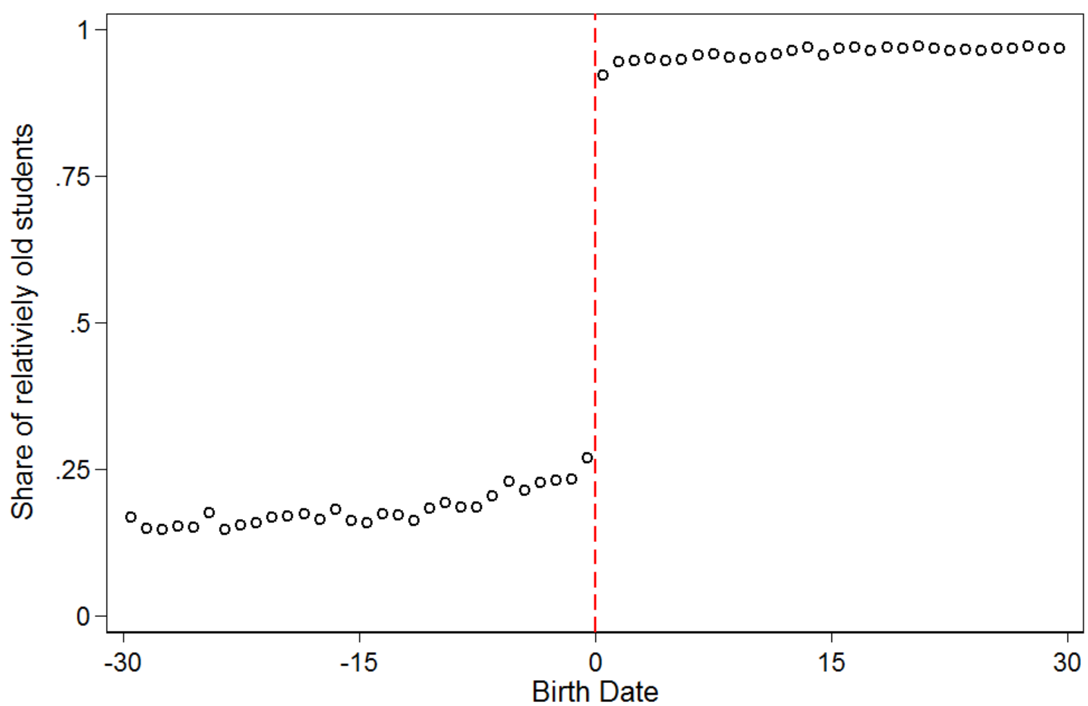

Figure 5.3: *The graph depicts how large share of students with a given birth date are classified as old in their class. In the horizontal axis runs the birth date, B_i, where zero is New Year. Author's rendering of data from Official Statistics of Finland.*

5.2. Main Results

First, I present graphical evidence for the relative age effect on special education. After that, I move on to regression tables and present the point estimates for different outcomes. Then, I present the heterogeneity results for different genders and educational backgrounds. Finally, I present results on the variation in time for the partially individualized (PIC).

Figure 5.4 shows the share of students with PIC for different birth dates around New Year. I have linear regression lines estimated before and after New Year with the specification of equation 9. As we can see, there is a noticeable jump at the cutoff from over 3% to slightly over 2%. It means that higher share of students born in December have PIC compared with students born in January.

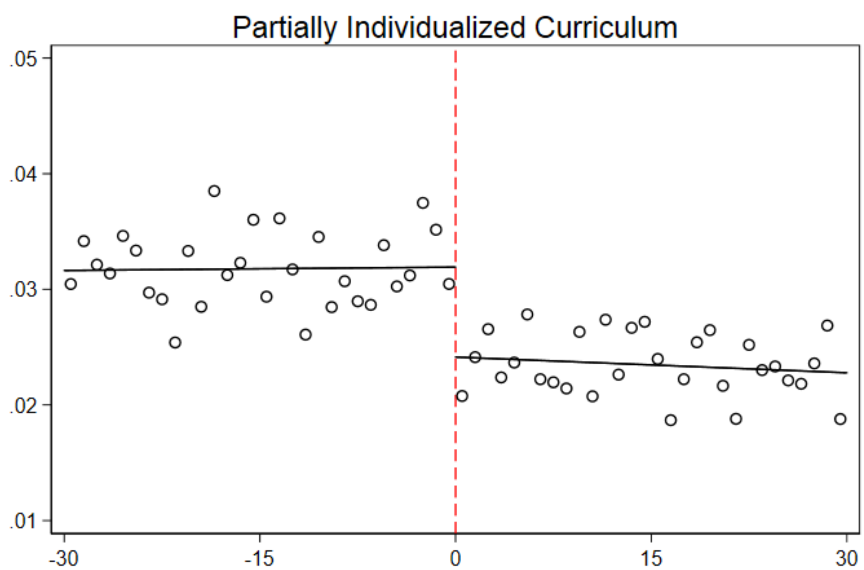

Figure 5.4: *The dots represent the share of students with partially individualized curriculum (PIC) with a given birth date. The lines are estimated using equation (7) a share of PIC as an outcome variable. The variable running on the horizontal axis is the birth date normalized around New Year. Author's rendering of data from Official Statistics of Finland.*

Similarly, Figure 5.5, which is on the same scale as Figure 5.4, shows the share of students with regular curriculum. The jump between students born in December and January is slightly smaller and in the opposite direction. The figure shows a jump from slightly over 94% to above 95% at the cutoff, which is New Year. A higher share of students who are born in January has regular curriculum once graduating from middle school compared with students born in December.

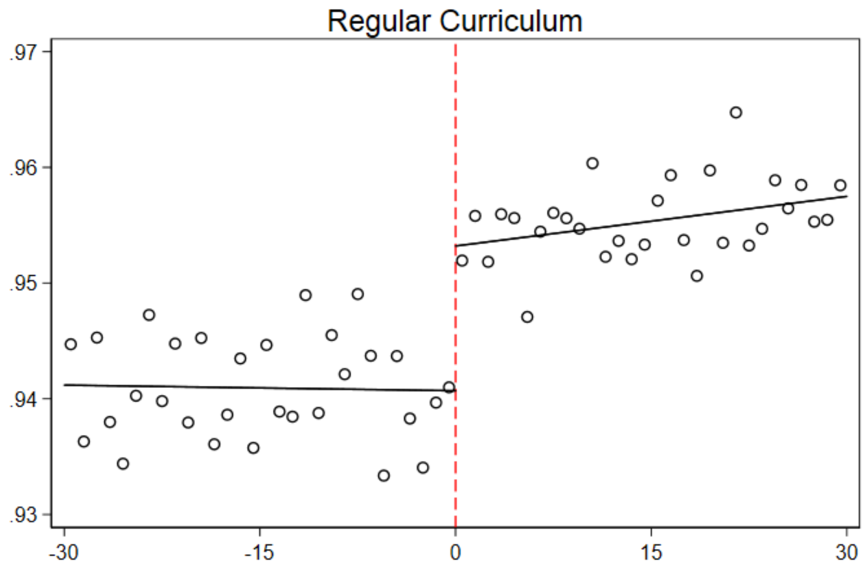

Figure 5.5: *The dots represent the share of students with regular curriculum with a given birth date. The lines are estimated using equation (7) a share of regular curriculum as an outcome variable. The variable running on the horizontal axis is the birth date normalized around New Year. Author's rendering of data from Official Statistics of Finland.*

In Figure 5.6, I show similar graph for students with MOFIC with the same scale as in Figure 5.4. As we can see, there is a jump at the cutoff but it is much smaller than in previous figures. A smaller jump was expected since students with MOFIC are much clearer cases and there is less room for misdiagnosis and subjectivity. Thus, relative school starting age and birth date should play smaller role in MOFIC.

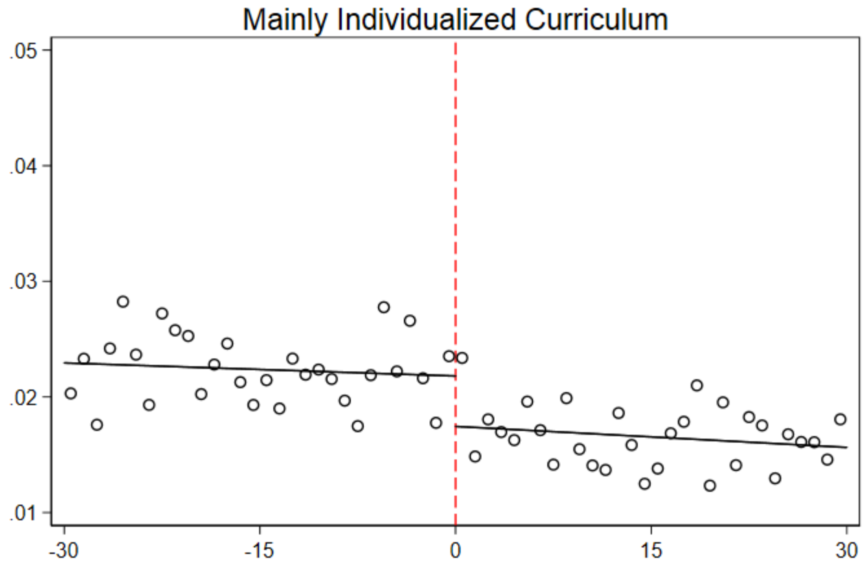

Figure 5.6: *The dots represent the share of students with mostly or fully individualized curriculum (MOFIC) with a given birth date. The lines are estimated using equation* (7) *with a share of MOFIC as an outcome variable. The variable running on the horizontal axis is the birth date normalized around New Year. Author's rendering of data from Official Statistics of Finland.*

In Table 5.3, I present the results of the PIC and see that younger students are more likely to graduate with a PIC. The rows of the first column present the regressions of equations (8), (7), and (9) respectively, with no control variables nor fixed effects. The second column presents estimation results for the same equations with control variables included and column (3) has both control variables and year fixed effects. In columns (1)–(3), the bandwidth is 30 days on each side of the cutoff, meaning the sample consists of the students who were born between December 2 and January 30. In column (4), I include 180 days of bandwidth on both sides. The specification includes the control variables, yearly fixed effects, and a linear function of the birth date. For all the estimation values, I present the standard errors which are clustered at a school level.

The reduced form estimates show that being one year older in relative age has an effect of -1.0 percentage points in the likelihood of having a regular curriculum. This result is robust for adding

Table 5.3: *Partially Individualized Curriculum*

	(1)	(2)	(3)	(4)
First Stage	0.698	0.715	0.715	0.795
s.e.	(0.0102)	(0.0111)	(0.0109)	(0.0108)
Reduced Form	-0.008	-0.010	-0.010	-0.009
s.e.	(0.0019)	(0.0021)	(0.0021)	(0.0008)
IV	-0.012	-0.013	-0.014	-0.011
s.e.	(0.0028)	(0.0030)	(0.0030)	(0.0010)
Control Variables	No	Yes	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
Bandwidth	30 days	30 days	30 days	180 days
Sample Size	157859	114139	114139	741285

Notes: Estimation results for partially individualized curriculum as an outcome. Standard errors, which are clustered at a school level, are in parentheses. For the first stage row I have used equation (8) for estimation, and, respectively, equations (7) and (9) for the reduced form and the IV.

control variables and yearly fixed effects. Estimation with a bandwidth of 180 days gives similar results even if the effect is slightly smaller, at -0.9 percentage points. Since Gelman and Imbens (2017) argue that RD effects with a high order of polynomials are usually not as accurate as linear or quadratic forms of the assignment variable, I show the 3rd degree polynomials in Appendix in Table 5.9.

As we can see from the first stage row, birth date estimates student's school starting age quite well. The estimate varies mostly between 0.7 and 0.8, which means that most students follow the school starting age rule. However, the value means that there are students who do not follow the rule and therefore the usage FRD design is justified. This fuzziness implies that the coefficient for the

IV is larger than in reduced form, as explained in subsection 3.2. As expected, we can see from the IV row that this is the case and based on the IV regressions, being one year older means that a student is -1.4 percentage points more likely to graduate with a PIC. This is a relatively large effect since on average only 2.8% of all students graduate with a PIC. Thus, one year in relative age reduces the probability of having a PIC by 46%. Moreover, this coefficient is robust for all the specifications.

As we can see from Table 5.3, the standard errors are quite small in all the estimations. Thus, we can be confident that relative age does have an effect on the probabilities of receiving special education. However, it is a more difficult to answer how large the effect is and how robust the estimates are. That is why I have multiple specifications to see whether the estimates vary on the specification or if they stay fairly constant.

The format of Table 5.4 is similar to Table 5.3 with only a different outcome variable. Here, I present the estimation results for the probability of graduating with a regular curriculum. Again, the standard errors are low for all the estimates. In reduced form the effect of relative age is estimated at around 1.3 percentage points. It means that one additional year in relative age reduces the probability of having a regular curriculum by 1.3 percentage points. For the IV setup this effect is estimated at 1.8 percentage points. The estimate remains at similar size with a larger bandwidth, as we can see from column (4).

Table 5.4: *Regular Curriculum*

	(1)	(2)	(3)	(4)
First Stage	0.698	0.715	0.715	0.795
s.e.	(0.0102)	(0.0111)	(0.0109)	(0.0108)
Reduced Form	0.013	0.013	0.013	0.014
s.e.	(0.0026)	(0.0027)	(0.0027)	(0.0011)
IV	0.019	0.018	0.018	0.017
s.e.	(0.0037)	(0.0037)	(0.0037)	(0.0014)
Control Variables	No	Yes	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
Bandwidth	30 days	30 days	30 days	180 days
Sample Size	159803	114921	114921	735633

Notes: Estimation results for regular curriculum as an outcome, standard errors, which are clustered at a school level, are in parentheses. Equation (8) is the first stage, and, respectively, equations (7) and (9) for the reduced form and the IV.

Next, I present the heterogeneity analysis to show how the relative age effect varies among different subsamples. First, I study the relative age effect for male and female graduates separately. Then, I study the effect for graduates whose neither parent has a higher education and for those who have at least one parent with higher education. I use the specification described in equations (8) and (9). I use the bandwidth of 30 days and I include control variables and year fixed effects.

In Table 5.5, I present the relative age effect on regular curriculum, PIC, and MOFIC for both male and female graduates. I include the MOFIC students also as comparison group, since relative age should not affect their diagnosis. As we can see from the first stage row, girls seem to follow the recommended school starting age more strictly than boys since first stage coefficient is larger for girls than boys. This is a statistically significant difference, which can be seen in Table A.4 in

Table 5.5: *Gender Differences*

	Regular		Partially Individualized		Mainly or Fully Individualized	
	Male	Female	Male	Female	Male	Female
First Stage	0.669	0.762	0.669	0.762	0.669	0.762
s.e.	(0.0116)	(0.0119)	(0.0116)	(0.0119)	(0.0116)	(0.0119)
Reduced Form	0.011	0.015	-0.006	-0.013	-0.004	-0.001
s.e.	(0.0040)	(0.0034)	(0.0031)	(0.0028)	(0.0024)	(0.0018)
IV	0.017	0.019	-0.010	-0.017	-0.006	-0.001
s.e.	(0.0060)	(0.0045)	(0.0047)	(0.0038)	(0.0035)	(0.0023)
Sample Size	58594	56327	58594	56327	58594	56327

Notes: Estimation results for both genders separately as an outcome. The standard errors are in parentheses, the bandwidth used is 30 days on both sides of the cutoff. The specifications are described in equations (8), (7). and (9).

Appendix. There is not much difference in the estimates for regular curriculum between female and male graduates, except for the fact that there is more variation for boys. This can be noticed from larger standard errors for boys. When it comes to PIC, the relative age effect for female graduates is -0.017 whereas for males it is only -0.010 and statistically significant only at a confidence level of 5%.

I consider graduates whose neither parent has a higher education to category Low and otherwise to category High. As we can see from the first stage row of Table 5.6, graduates with high education background are less likely to follow the school starting age rule. However, as presented in Appendix, the difference between first stages is not statistically significant for different educational backgrounds. We can see that the relative age effect is stronger for the low education background in both regular curriculum and PIC. For graduates with a high education background, the relative age effect for the PIC is only significant at a confidence level of 5%. However, there is a concern that

Table 5.6: *Educational Background Differences*

	Regular		Partially Individualized		Mainly or Fully Individualized	
Parents' Education:	Low	High	Low	High	Low	High
First Stage	0.716	0.679	0.716	0.679	0.716	0.679
s.e.	(0.0136)	(0.0086)	(0.0136)	(0.0086)	(0.0136)	(0.0086)
Reduced Form	0.017	0.012	-0.013	-0.006	-0.003	-0.006
s.e.	(0.0035)	(0.0038)	(0.0028)	(0.0027)	(0.0020)	(0.0024)
IV	0.023	0.017	-0.017	-0.008	-0.004	-0.009
s.e.	(0.0049)	(0.0056)	(0.0040)	(0.0039)	(0.0028)	(0.0036)
Sample Size	80282	77530	80282	77530	80282	77530

Notes: Estimation results for middle school graduates whose neither parent has a higher education compared (Low) and whose at least one parent has higher education (High). The bandwidth used is 30 days on both sides of the cutoff. The specifications are described in equations (8), (7). and (9).

the relative age effects are quite large for MOFIC students, especially with higher parental education background. This is unexpected, since assigning to MOFIC is considered as a very objective process.

Finally, I introduce the results of the yearly analysis. Figure 5.7 presents the FRD estimates for PIC graduates each year separately. The estimation equation included all the control variables and had a bandwidth of 30 days a side which is the specification of Table 5.3, column (2). As we can see from the figure, the confidence intervals of the estimates are relatively large and there are not many statistically significant point estimate with a 95% confidence. However, we can see a downward trend in the estimates from 2005 onwards. There is a statistically significant negative relative age effect for 2008 and 2012–2014.

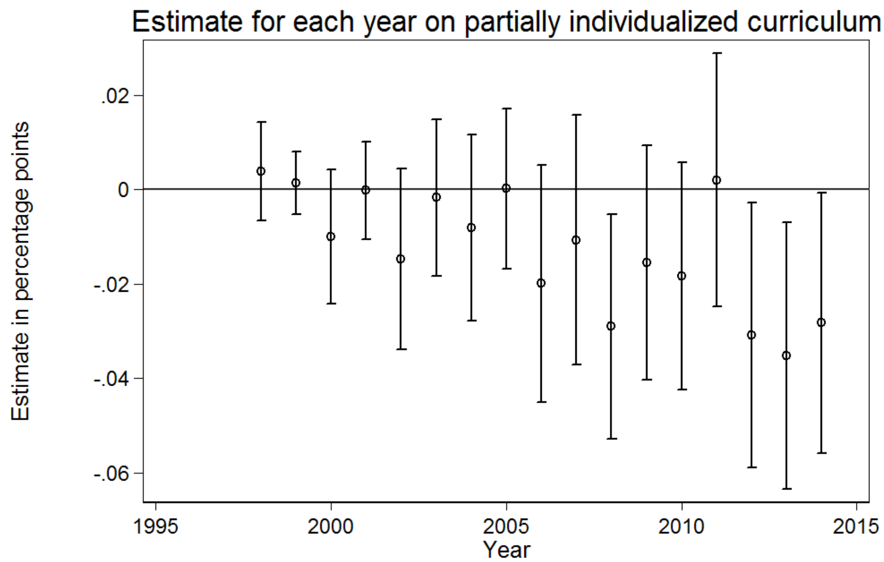

Figure 5.7: *The relative age effect for each year separately on partially individualized curriculum. The regressions included all the control variables and a bandwidth of 30 days (as in Table 5.3, column (2)). The point estimates are from the IV specification.*

In Figure 5.8, I have used a larger sample of 360 days to improve the precision of the estimates. As we can see, the confidence intervals are smaller and the trend is similar to the one in Figure 5.7. The red diamonds in the figure represent the point estimates obtained from Figure 5.7. The figure shows that the point estimates varied depending on the chosen bandwidth. Mostly, the estimation with a larger bandwidth gives smaller point estimates in absolute terms. To emphasize this change in the trend of relative age effect, I run an FRD regression before and after the year 2005. Table A.5 in Appendix confirms the interpretation of the graph that there is no relative age effect before the year 2005 and a strong effect after that.

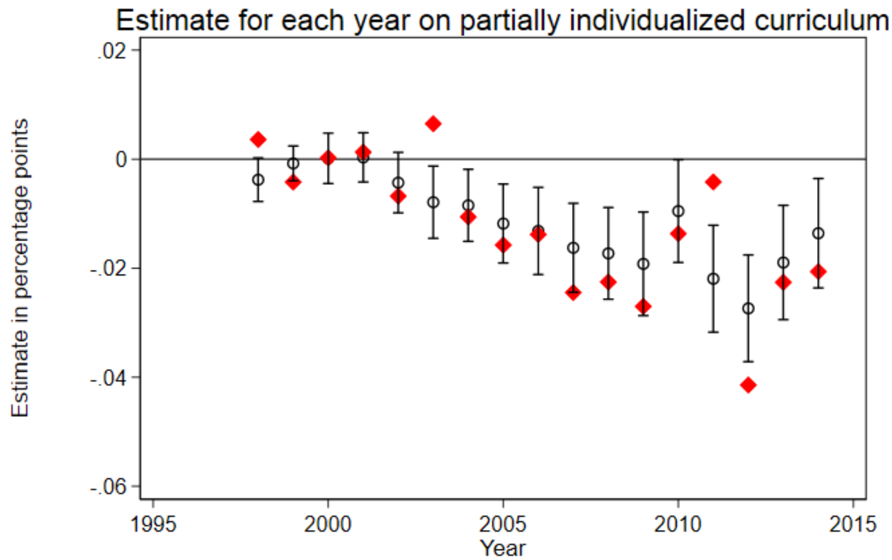

Figure 5.8: *The relative age effect for each year separately. The regressions included all the control variables and a bandwidth of 360 days. The red diamonds are the point estimates of Figure 5.7. The point estimates are from the IV specification. Author's rendering of data from Official Statistics of Finland.*

5.3. Robustness

In this subsection, I do robustness checks to test whether the estimates vary with different specifications. First, I show graphs that present the point estimates with different bandwidth choices. Then, I show placebo tests for the outcomes by setting the cutoff point to different values than New Year. Finally, I show the estimation results with uniform weights. I show robustness tests only for regular curriculum and PIC since there is no consistent effect found for MOFIC. However, similar figures for MOFIC are in Appendix.

Even if there are robust ways of choosing a bandwidth, there is always some arbitrariness in its selection (see, for example, Imbens and Lemieux, 2008). Therefore, it is important to show that the results are not dependent on the chosen bandwidth. I use a graph to show how the point estimates and confidence intervals (or standard errors) change with different bandwidths. The specification

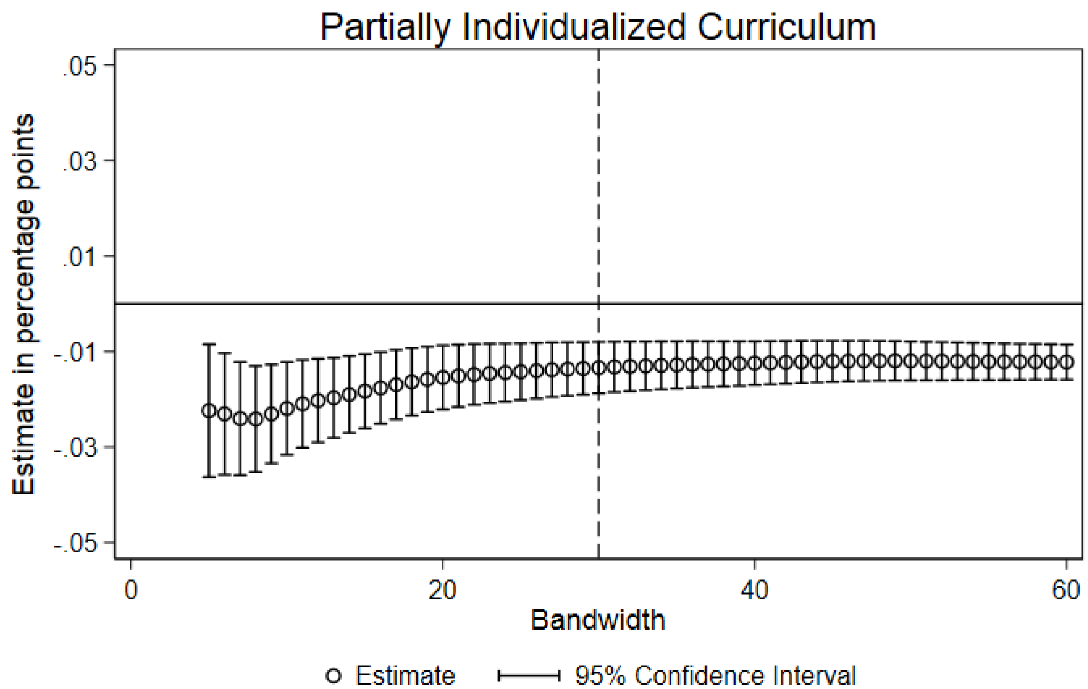

Figure 5.9: *Estimation points and 95% confidence intervals for different bandwidths for partially individualized curriculum. The horizontal axis is the length of bandwidth on one side of the cutoff. Author's rendering of data from Official Statistics of Finland.*

in the graphs is the same as before with all the controls, but I run the regressions with every bandwidth from 5 to 60 on each side.

Figure 5.9 shows how the point estimate and confidence interval vary with different bandwidth choice for PIC. The point estimate with a bandwidth of 5 days per side is around -0.02 and when the bandwidth increases it converges to slightly below -0.01. Respectively, the confidence interval decreases when the bandwidth widens. Thus, we can conclude that the point estimate varies slightly with different bandwidths, but it has a statistically significant value.

In Figure 5.10, I present a similar graph for the effect on the regular curriculum. The trend is very similar to the one in the previous figure, except for the fact that the values are positive. The point

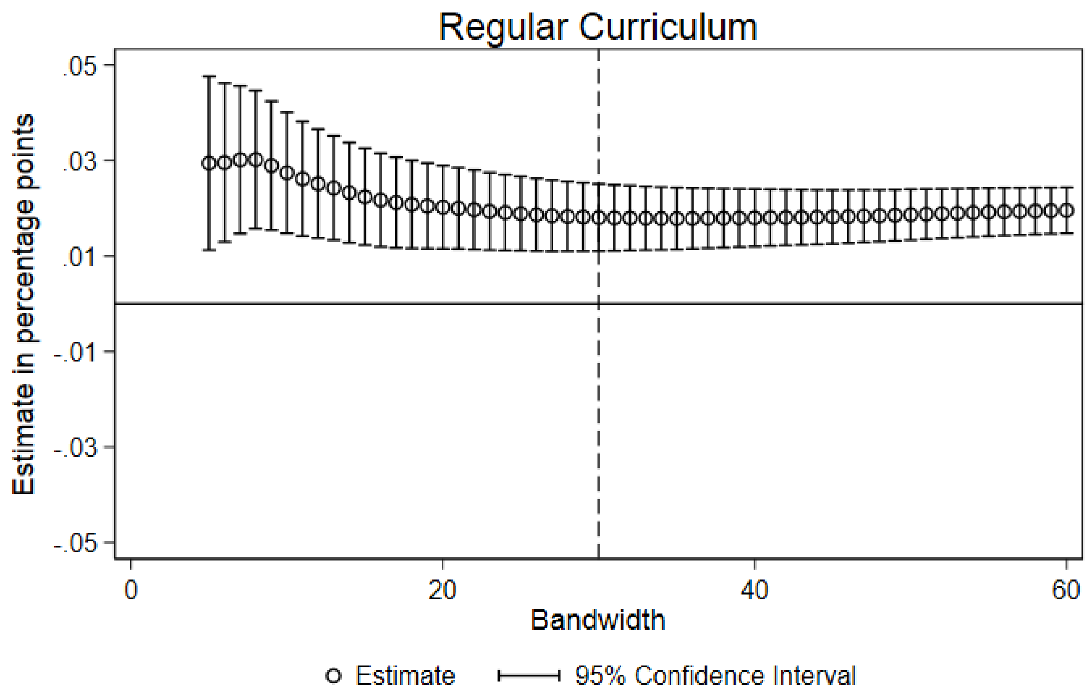

Figure 5.10: *Estimation points and 95% confidence intervals for different bandwidths for regular curriculum. The horizontal axis is the length of bandwidth on one side of the cutoff. Author's rendering of data from Official Statistics of Finland.*

estimate with a bandwidth of 5 days per side is around 0.03 and it converges towards slightly below 0.02 as the bandwidth increases. The confidence interval decreases and the point estimate stays at a clearly statistically significant level.

Next, I will show graphs that present placebo tests. As suggested by Imbens and Lemieux (2008), I run regressions for different cutoffs that are non-discontinuity points. With this analysis, I pursue to convince that there really is an effect at a discontinuity point and it does not appear only by chance. The following three graphs include the regression point estimates from equations (8) and (9) with a bandwidth of 30 days a side. Cutoff points are every fifth date from -25 to 25 (where 0 is New Year).

In Figure 5.11, I have regression estimates and 95% confidence intervals for the PIC at nine different cutoffs. As we can see, the only statistically significant effect appears at 0 which is New Year. With this cutoff, the effect is -1.5 percentage points whereas other point estimates are close to zero. Thus, we can be fairly confident that the effect does not appear by random chance. The discontinuity in assignment variable seems to cause an effect on the outcome.

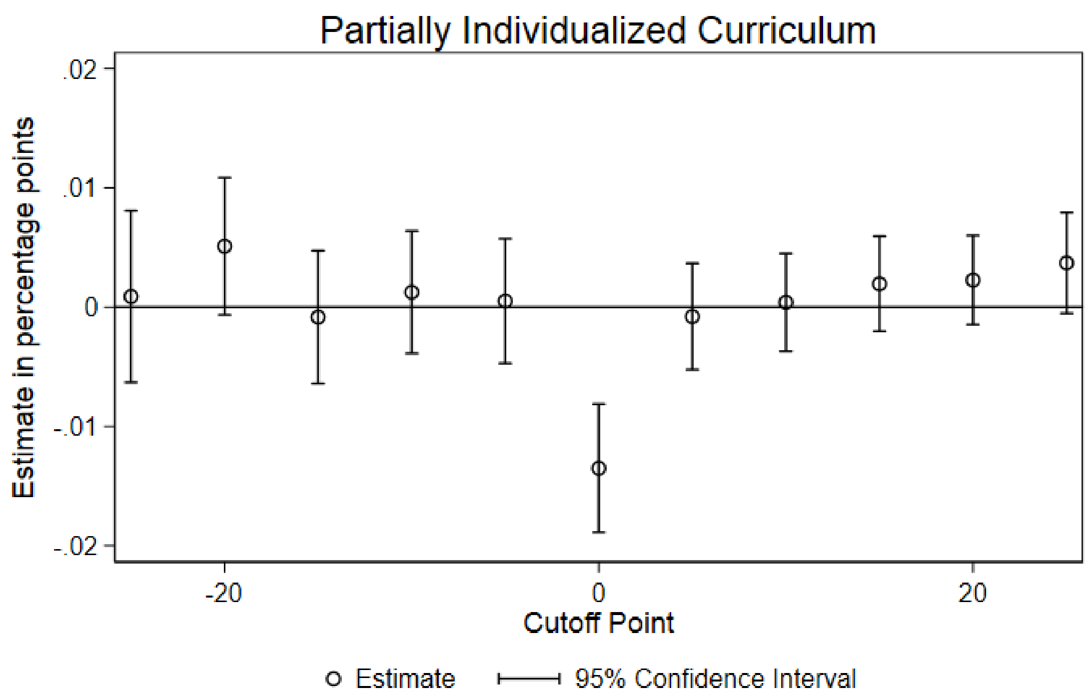

Figure 5.11: *Estimation points and 95% confidence intervals for placebo cutoffs for partially individualized curriculum. The horizontal axis shows the cutoff date used in the estimation. Author's rendering of data from Official Statistics of Finland.*

Similarly, for the regular curriculum, we can note from Figure 5.12, that the only statistically significant effect is at a cutoff point 0 again. The point estimate for that cutoff point is almost two percentage points whereas the point estimates for all the other cutoff points lies around half a percentage points from zero.

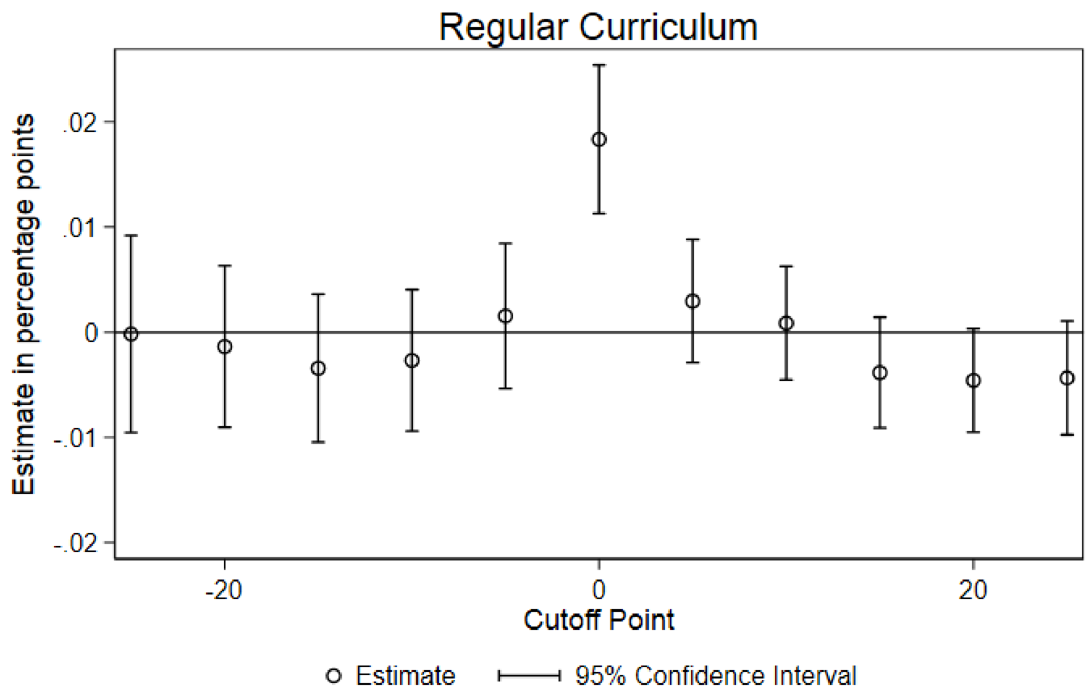

Figure 5.12: *Estimation points and 95% confidence intervals for placebo cutoffs for regular curriculum. The horizontal axis shows the cutoff date used in the estimation. Author's rendering of data from Official Statistics of Finland.*

In Tables 5.7 and 5.8, I present that the estimates are insensitive for the choice of kernel function. Table 5.7 shows otherwise same estimation results as in Table 5.3 but with a uniform kernel function. This means that every observation has an equal weight regardless of its distance from the cutoff. Table 5.7 shows that the estimates are almost constant for all specifications with -0.8 percentage points in reduced form and from -1.0 to -1.1 in the IV specification. These are very similar to the estimates we obtained with a triangular kernel function.

Table 5.7: *Robustness Test for the Kernel Function for PIC*

	(1)	(2)	(3)	(4)
First Stage	0.712	0.729	0.730	0.816
s.e.	(0.0101)	(0.0107)	(0.0105)	(0.0111)
Reduced Form	-0.008	-0.008	-0.008	-0.008
s.e.	(0.0018)	(0.0019)	(0.0019)	(0.0007)
IV	-0.011	-0.011	-0.011	-0.010
s.e.	(0.0025)	(0.0026)	(0.0026)	(0.0009)
Control Variables	No	Yes	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
Bandwidth	30	30	30	180
Sample Size	159803	114921	114921	735633

Notes: A robustness analysis for the choice of kernel function. In the main results, I used a triangular kernel function but the estimations for this table is obtained with uniform kernel function. The outcome variable is PIC, school level clustered standard errors are in the parentheses. The specifications are equation (8) for first stage, equation (7) for reduced form and equation (9) for the IV.

Table 5.8 shows the same estimates for regular curriculum. Again, the estimates do not vary much with different specifications. The reduced form estimates are around 1.2 and 1.3 percentage points and the IV is between 1.5 and 1.8 percentage points. Again, these are only slightly smaller estimates than in Table 5.4. With this analysis we can conclude that the estimates are not sensitive to the choice of kernel function.

Table 5.8: *Robustness Test for the Kernel Function for Regular Curriculum*

	(1)	(2)	(3)	(4)
First Stage	0.712	0.729	0.730	0.816
s.e.	(0.0101)	(0.0107)	(0.0105)	(0.0111)
Reduced Form	0.013	0.012	0.012	0.013
s.e.	(0.0024)	(0.0024)	(0.0024)	(0.0010)
IV	0.018	0.017	0.017	0.015
s.e.	(0.0033)	(0.0033)	(0.0033)	(0.0013)
Control Variables	No	Yes	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
Bandwidth	30	30	30	180
Sample Size	159803	114921	114921	735633

Notes: A robustness analysis for the choice of kernel function. In the main results, I used a triangular kernel function but the estimations for this table is obtained with uniform kernel function. The outcome variable is regular curriculum, school level clustered standard errors are in the parentheses. The specifications are equation (8) for first stage, equation (7) for reduced form and equation (9) for the IV.

Finally, I study whether the estimations are sensitive to the choice of the degree of the polynomial function in the estimation equations. In the main results, I use linear estimations but in some cases increasing the degree of the polynomial functions can increase the accuracy of the estimations. However, as Gelman and Imbens (2017) argue, usually the linear and quadratic forms provide the best accuracy. Thus, in Table 5.9, I present the estimation results with the second order polynomials in equations (7), (8), and (9). As we can see, the estimation results do not change much with the quadratic form. For both the reduced form and IV, there is a change of 0.2 percentage points at most, when we compare these results to the corresponding estimates in Tables 5.3 and 5.4. Thus, the results are not sensitive to the choice of the degree of the polynomial function in the estimation equations.

Table 5.9: *Regressions with Quadratic Fom*

Outcome	PIC		Regular Curriculum	
	(1)	(2)	(3)	(4)
First Stage	0.672	0.732	0.672	0.732
s.e.	(0.0106)	(0.0106)	(0.0106)	(0.0106)
Reduced Form	-0.009	-0.009	0.013	0.014
s.e.	(0.0023)	(0.0009)	(0.0031)	(0.0013)
IV	-0.014	-0.012	0.021	0.021
s.e.	(0.0041)	(0.0015)	(0.0054)	(0.0021)
Bandwidth	30	180	30	180
Sample Size	159803	1019436	159803	1019436

Notes: Estimations for the PIC and regular curriculum with second order polynomial functions in the estimation equations (8), (7), and (9). There are no control variables nor fixed effects. School level clustered standard errors are in the parentheses.

6. Discussion

Based on the results of this thesis, relative school starting age affects the probabilities of having different types of individualization in the curriculum. My results are consistent with earlier studies which show that relatively younger children are more likely to be in special education, have an ADHD diagnosis, and do worse at school (Dhuey and Lipscomb, 2010; Elder and Lubotsky, 2009; Elder, 2010). To sum up my results, relatively older students are more likely to follow the regular curriculum and younger students are more likely to follow a partially individualized curriculum (PIC). Since the purpose of special education is to improve student's educational outcomes, it is important to study why children are assigned to special education. The question is whether younger students truly need more individualization and special care for their primary education. The other possibility that would explain the results is that younger children are more easily misplaced in special education due to their immaturity compared to older classmates. However, to answer this question is beyond the reach of this thesis. Since the data only contain information on the level of individualization at the end of compulsory school, it is not possible to see when children have started an individualized curriculum. This information would be interesting since relative age differences are larger at younger ages.

Another result of this thesis is that the relative age effect is stronger for girls on having a PIC. As we showed in the analysis, girls are more likely to follow the school starting age rule than boys. This could mean that boys who are born at the end of the year are more likely to be held back and start school a year later. Children who are held back may have disabilities related to learning with a higher probability. Therefore, a higher probability of being held back leads to a lower relative age effect.

The probability of being held back is greater for children with a higher educational background. The relative age effect is stronger for children whose neither parent has tertiary education. The relative age effect of one year for the PIC is more than double the size for children with low educational background compared to children with high educational background. The difference in the first stage of the regression is larger between genders than parents' educational backgrounds. Since the first stage of the regression tells about the compliance of the school starting age rule, gender has a larger impact on following the rule than parent's educational background. Thus, the difference in relative age effect between educational backgrounds probably arises from something else than different probabilities of following the school starting rule.

The third main result of the study is the change in trends over time. Restrictions on the data make it difficult to study timings of these trends since we only see the endpoint of students' compulsory education. Most of the special education decisions are done at latest in grade 7, so there is always a lag between the change in trend and when it emerges in the data. As explained in the previous section, the relative age effect on PIC was close to zero until 2005. After that, the relative age effect has been negative, meaning that younger children have been more likely to graduate with a PIC. Before 2005, we cannot reject the null hypothesis of the relative age effect on the PIC. In contrast, a year younger students in relative age have been 2 percentage points more likely to have a PIC on average since 2005. Thus, the change in trends of assigning individualized curricula has most likely happened sometime before 2005.

I do not find a change in trends from the special education and funding reforms of 2009 and 2010. As discussed above, it is possible that there is a lag in the impact of the reforms. One may have to accumulate more data to estimate their impact. Another explanation is that the reforms have not changed the relative age effect on PICs even if it has decreased the total number of special education students.

As stated in the literature review, there are some possible explanations for the trend changes. One of the most convincing ones is that special education became full inclusive in a classroom setup in the late 1990s. Thus, it increased the number of special education overall but especially in categories like “other reasons”. This increase may have led to more subjective decision making and thus, relative age may have had a larger impact on those decisions. In short, an increase in the supply of special education may have lead to an increased relative age effect. However, I need to emphasize that this is only a hypothesis that should be studied in more depth.

7. Conclusion

In this thesis, I studiy how relative school starting age affects the probability of having an individualized curriculum. With a bandwidth of 30 days on both sides of the cutoff, I find that students who are born in right after the school starting cutoff are 1.3 percentage points more likely to follow a regular curriculum than students born right before New Year. With fuzzy regression discontinuity the effect on compliers is 1.8 percentage points. Similarly, the effect of having a partially individualized curriculum is -1.0 percentage points in reduced form and -1.4 percentage points with FRD. The FRD estimate is equivalent to a 40% decrease in the probability since only 2.6% of all students follow partially individualized curriculum on average. These effects are significant and robust for multiple variations and robustness checks. For the mainly or fully individualized curriculum, I find a small negative effect, meaning that older students are less likely to graduate with it. However, this effect is not robust and the standard errors are quite large.

I study two heterogeneity specifications: gender and educational background. I find that the relative age effect is stronger for girls and students with a low educational background. However, the role of education is not as strong and consistent as gender. Even if it not possible to find the cause for these differences, the share of compliers is larger for girls and children with low education background. This can mean that they are more likely to follow the school starting age rule. I hypothesize that boys and children with high educational background are more likely to be held back. This would lead to larger selection bias in the complier group and, hence, the relative age effect is smaller for them.

The third result of my thesis is that the relative age effect for the partially individual curriculum emerged around the year 2005. It has been negative until 2014 for which I have the latest observations. Due to restrictions on the data, I am not able to find the cause of this change and it leaves space for further research. However, I hypothesize that the reason may reach back to the late 1990s when special education became fully inclusive and to the Core Curriculum Reform in 2004. I do not find any change in the effect occurring from the special education funding reform of 2009 and the Basic Education Act reform introduced in 2010. The effects of these reforms may not show up in my data which has the latest cohort graduating in 2014. Therefore, it is important to study these trends in the future. Once we have more data available, we can see whether the recent reforms have affected the relative age effect.

References

Ahtiainen, Raisa, "Reformin Implementaatio [The Implementation of a Reform]," *MSc Thesis, University of Helsinki*, 2010.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An empiricist's companion*, 1 ed., Princeton, New Jersey: Princeton University Press, 2009.

Barak, Yoram, Aliza Ring, Juliet Sulkes, Uri Gabbay, and Avner Elizur, "Season of Birth and Autistic Disorders in Israel," *The American Journal of Psychiatry*, 1995, *152* (5), 798–800.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Too Young to Leave the Nest? The Effects of School Starting Age," *Review of Economics and Statistics*, 2011, *93* (2), 455–467.

Cohen, Jessica L., "Causes and Consequences of Special Education Placement: Evidence from Chicago Public Schools," *Brookings Institution Working Paper*, 2007.

De Haan, Monique, "The Effect of Additional Funds for Low-ability Pupils: A Non-parametric Bounds Analysis," *Economic Journal*, 2017, *127* (599), 177–198.

Dhuey, Elizabeth and Stephen Lipscomb, "Disabled or Young? Relative Age and Special Education Diagnoses in Schools," *Economics of Education Review*, 2010, *29* (5), 857–872.

_ **and** _ , "Funding Special Education by Total District Enrollment: Advantages, Disadvantages, and Policy Considerations," *Education Finance and Policy*, 2013, *8* (3), 316–331.

Dickert-Conlin, Stacy and Todd Elder, "Suburban Legend: School Cutoff Dates and the Timing of Births," *Economics of Education Review*, 2010, *29* (5), 826–841.

Elder, Todd E., "The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates," *Journal of Health Economics*, 2010, *29* (5), 641–656.

— **and Darren H. Lubotsky**, "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, 2009, *44* (3), 641–683.

Fan, Jianqing and Irène Gijbels, "Variable Bandwidth and Local Linear Regression Smoothers," *The Annals of Statistics*, 1992, *20* (4), 2008–2036.

— **, Theo Gasser, Irène Gijbels, Michael Brockmann, and Joachim Engel**, "Local Polynomial Regression: Optimal Kernels and Asymptotic Minimax Efficiency," *Annals of the Institute of Statistical Mathematics*, 1997, *49* (1), 79–99.

Finnish Basic Education Act, 628/1998.

Finnish National Agency for Education, "Perusopetuksen Vuoden 2004 Opetussuunnitelmauudistus – Kehittämisverkostoon ja kokeiluun osallistuneiden kuntien ja koulujen näkemyksiä ja ratkaisuja [The Core Curriculum Reform of 2004 – Views and solutions of the municipalities and schools that participated in the development network and the experiment]," 2007.

— , "LATU – Erityistä Tukea Tarvitsevien Opetuksen Kehittäminen Yleis- ja Erityisopetuksessa 2002–2004 [LATU – The Development of Special Needs Education in General and Special Education 2002–2004]," September 2017.

Fredriksson, Peter and Björn Öckert, "Life-cycle Effects of Age at School Start," *The Economic Journal*, 2014, *124* (579), 977–1004.

Gans, Joshua S. and Andrew Leigh, "Born on the First of July: An (un)natural Experiment in Birth Timing," *Journal of Public Economics*, 2009, *93* (1-2), 246–263.

Gelman, Andrew and Guido W. Imbens, "Why High-order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2017, (accepted).

Graham, Linda J. and Markku Jahnukainen, "Wherefore Art Thou, Inclusion? Analysing the Development of Inclusive Education in New South Wales, Alberta and Finland," *Journal of Education Policy*, 2011, *26* (2), 263–288.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin, "Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities?," *The Review of Economics and Statistics*, 2002, *84* (4), 584–599.

Heckman, James J., "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 2006, *312* (5782), 1900–1902.

Henry, Gary T., C. Kevin Fortner, and Charles L. Thompson, "Targeted Funding for Educationally Disadvantaged Students: A Regression Discontinuity Estimate of the Impact on High School Student Achievement," *Educational Evaluation and Policy Analysis*, 2010, *32* (2), 183–204.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

Iversen, Jon Marius Vaag, "School Accountability Reforms and the Use of Special Education," *Education Economics*, 2013, *21* (3), 264–280.

Kaila, Martti, "The Effects of Relative School Starting Age on Educational Outcomes in Finland," *VATT Working Papers*, 2017, *84*, 1–60.

Keslair, Francois, Eric Maurin, and Sandra McNally, "Every Child Matters? An Evaluation of "Special Educational Needs" Programmes in England," *Economics of Education Review*, 2012, *31* (6), 932–948.

Kirjavainen, Tanja, Jonna Pulkkinen, and Markku Jahnukainen, "Erityisoppilaiden Osuuk-sien Kuntakohtaiseen Vaihteluun Vaikuttaneet Tekijät vuosina 2001–2010 [The Factors of the Changes in the Shares of Special Education Students at a Municipality-level in Years 2001–2010],” *Yhteiskuntapolitiikka*, 2014, *79* (6), 619–630.

_ , _ , and _ , “Special Education Students in Transition to Further Education: A Four-year Register-based follow-up Study in Finland,” *Learning and Individual Differences*, 2016, *45*, 33–42.

Kwak, Sally, “The Impact of Intergovernmental Incentives on Student Disability Rates,” *Public Finance Review*, 2010, *38* (1), 41–73.

Lahtinen, Matti and Timo Lankinen, *Koulutuksen Lainsäädäntö Käytännössä [The Law of Education in Practice]*, 9 ed., Helsinki, Finland: Tietosanoma OY, 2015.

Lavy, Victor and Analía Schlosser, “Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits,” *Journal of Labor Economics*, 2005, *23* (4), 839–874.

Lee, David S. and Thomas Lemieux, “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature*, 2010, *48* (2), 281–355.

McCrary, Justin, “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test,” *Journal of Econometrics*, 2008, *142*, 698–714.

Ministry of Education, “Erityisopetuksen Strategia [Special Education Strategy],” *Reports of the Ministry of Education, Finland*, 2007, *47.*

National Audit Office of Finland, “Erityisopetus Perusopetuksessa [Special Needs Education in Basic Education],” 2013.

Naukkarinen, Aimo, “Laatua Erityisopetukseen vai Yleisopetukseen? [Quality for the Special Education or the General Education?],” *Finnish National Agency for Education*, 2005.

Official Statistics of Finland, “Special Education [e-publication],” 2016.

Pulkkinen, Jonna and Markku Jahnukainen, “Finnish Reform of the Funding and Provision of Special Education: The Views of Principals and Municipal Education Administrators,” *Educational Review*, 2016, *68* (2), 171–188.

Rinkinen, Aija and Minna Lindberg, “Tuen Portailla – Tehostetun ja erityisen tuen kehittämistoiminta 2008–2012 [The Stairs of Support – The development of intensified and special support 2008–2012],” *Finnish National Agency for Education*, 2014.

Sayal, Kapil, Roshan Chudal, Susanna Hinkka-Yli-Salomäki, Petteri Joelsson, and Andre Sourander, “Relative Age Within the School Year and Diagnosis of Attention-Deficit Hyperactivity Disorder: A Nationwide Population-Based Study,” *The Lancet Psychiatry*, 2017, *4* (11), 868–875.

Schwandt, Hannes and Amelie Wuppermann, “The Youngest Get the Pill: ADHD Misdiagnosis in Germany, Its Regional Correlates and International Comparison,” *Labour Economics*, 2016, *43*, 72–86.

Setren, Elizabeth, “Special Education and English Language Learner Students in Boston Charter Schools: Impact and Classification,” *School Effectiveness and Inequality Initiative Discussion Paper*, 2016, *4*, 1–68.

Silliman, Mikko I., “Targeted Funding, Immigrant Background, and Educational Outcomes: Evidence from Helsinki's "Positive Discrimination" Policy,” *VATT Working Papers*, 2017, *91*, 1–49.

Thuneberg, H., J. Hautamäki, R. Ahtiainen, M. Lintuvuori, M.-P. Vainikainen, and T. Hilasvuori, “Conceptual Change in Adopting the Nationwide Special Education Strategy in Finland,” *Journal of Educational Change*, 2014, *15* (1), 37–56.

A. Appendix

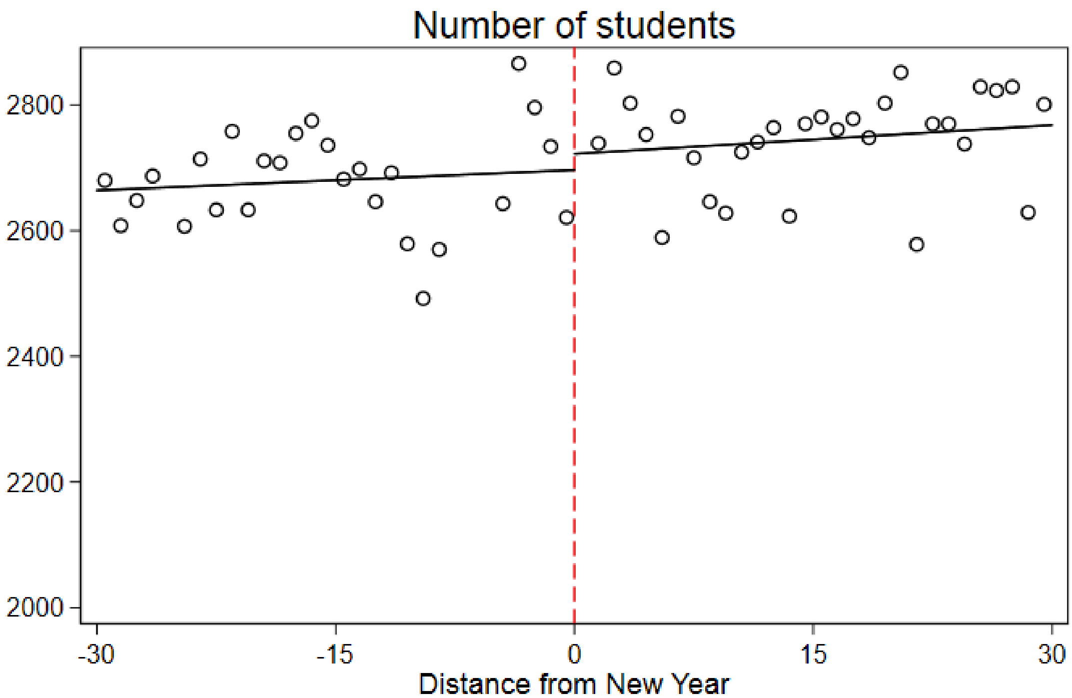

Figure A.1: *This presents the same graph as Figure 5.1 with deleted outliers. The lines are local linear regressions on both sides of New Year for the number of students. Outliers were the Christmas dates (Dec 24–26), Independence day (Dec 6) and New Year (Jan 1). Author's rendering of data from Official Statistics of Finland.*

Table A.1: *Share of Students with Individualization*

Year	9th Graders in Special Support	Any Individualized Curriculum	Share of individualized in Special Support
1998	2477	1429	0,58
1999	2737	1489	0,54
2000	3034	1658	0,55
2001	3454	1773	0,51
2002	3811	1865	0,49
2003	4187	2273	0,54
2004	4548	2522	0,55
2005	5229	2700	0,52
2006	5602	3425	0,61
2007	6003	3527	0,59
2008	5979	3742	0,63
2009	6106	4039	0,66
2010	5997	3946	0,66
2011	6886	3899	0,57
2012	6222	3877	0,62
2013	5979	3403	0,57
2014	5759	3546	0,62
In total:	84010	49113	0,58

Notes: The table demonstrates the share of students with individualized curriculum in special support. Every student with an individualization are in special support but some special support students follow regular curriculum. The second column shows the number of 9th graders in special support in Finland for a given year. The third column is a sum of middle school graduates with different levels of individualization for each year. These two values are obtained from different data sets and, thus, the last column should be interpreted as an estimation of the share of students with an individualized curriculum in special support for each year.

Table A.2: *Comparison of Samples*

	RD Sample	Excluded Sample	Difference	P-value
Mother's Education	4.23	4.23	0.01	0.07
Mother's Earnings	25,577.96	25,510.10	67.86	0.55
Father's Education	4.25	4.24	0.01	0.06
Father's Earnings	28,819.30	28,893.79	-74.48	0.63
GPA	7.61	7.61	-0.00	0.61
Regular Curriculum	0.95	0.95	0.00	0.70
Partially Individualized Curriculum	0.03	0.03	0.00	0.79
Mainly or Fully Individualized Curriculum	0.02	0.02	-0.00	0.90
Sample size	155,841	852,717	In Total:	1,008,558.00

Notes: Comparison between the RD sample and the excluded sample. In the RD sample, I have all students who were born in 30 days on both sides of the cutoff. In the second and third column, I present the averages of the control and outcome variables with both samples. In the fourth column, there is the difference of these averages. The p-value of these differences are presented in the final column. In principle, the p-values present how likely it is to obtain these two samples from the same distribution.

Table A.3: *Mostly or Fully Individualized Curriculum*

	(1)	(2)	(3)	(5)
First Stage	0.698	0.715	0.715	0.795
s.e.	(0.0102)	(0.0111)	(0.0109)	(0.0108)
Reduced Form	-0.004	-0.002	-0.002	-0.004
s.e.	(0.0015)	(0.0015)	(0.0015)	(0.0007)
IV	-0.006	-0.003	-0.003	-0.005
s.e.	(0.0022)	(0.0020)	(0.0021)	(0.0009)
Control Variables	No	Yes	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
Bandwidth	30 days	30 days	30 days	180 days
Polynomial Order	1st	1st	1st	1st
Sample Size	157859	114139	114139	730481

Notes: Estimation results for mainly or fully individualized curriculum as an outcome. School level clustered standard errors are in parentheses. For the first stage row I have used equation (8) for estimation, and, respectively, equations (7) and (9) for the reduced form and the IV.

Table A.4: *The Effect of Gender and Educational Background*

	First Stage	Regular	Partially Individualized	Mainly or Fully Individualized
First Stage	0.699			
s.e.	(0.0101)			
IV		0.019	-0.012	-0.006
s.e.		(0.0037)	(0.0028)	(0.0022)
Female	-0.081	0.025	-0.012	-0.013
s.e.	(0.0027)	(0.0018)	(0.0010)	(0.0012)
High education	-0.012	-0.002	-0.005	0.005
s.e.	(0.0031)	(0.0016)	(0.0011)	(0.0010)
Sample Size	159803	159803	159803	159803

Notes: The role of gender and family education background on the assignment variable and outcome variables. Both gender and education background have a significant effect on the first stage, meaning that there is a difference in following the school starting age rule between these subgroups. Especially gender has a statistically significant and large effect on all outcomes. The specification of the regressions are presented in equations (8) and (9) and the bandwidth is 30 days on both sides of the cutoff.

Table A.5: *Effect Before and After 2005.*

		Regular		Partially Individualized		Mainly or Fully Individualized	
Before	IV	0.005	0.008	-0.003	-0.004	-0.001	-0.003
	s.e.	(0.005)	(0.002)	(0.003)	(0.001)	(0.003)	(0.001)
	Sample Size	47 403	300 780	47 403	300 780	47 403	300 780
After	IV	0.027	0.024	-0.021	-0.017	-0.005	-0.007
	s.e.	(0.005)	(0.002)	(0.005)	(0.002)	(0.003)	(0.001)
	Sample Size	67 518	434 853	67 518	434 853	67 518	434 853
	Bandwidth	30	180	30	180	30	180

Notes: Estimation results in FRD separately before and after 2005, standard errors are in parentheses. I use the estimation equations (8) and (9).

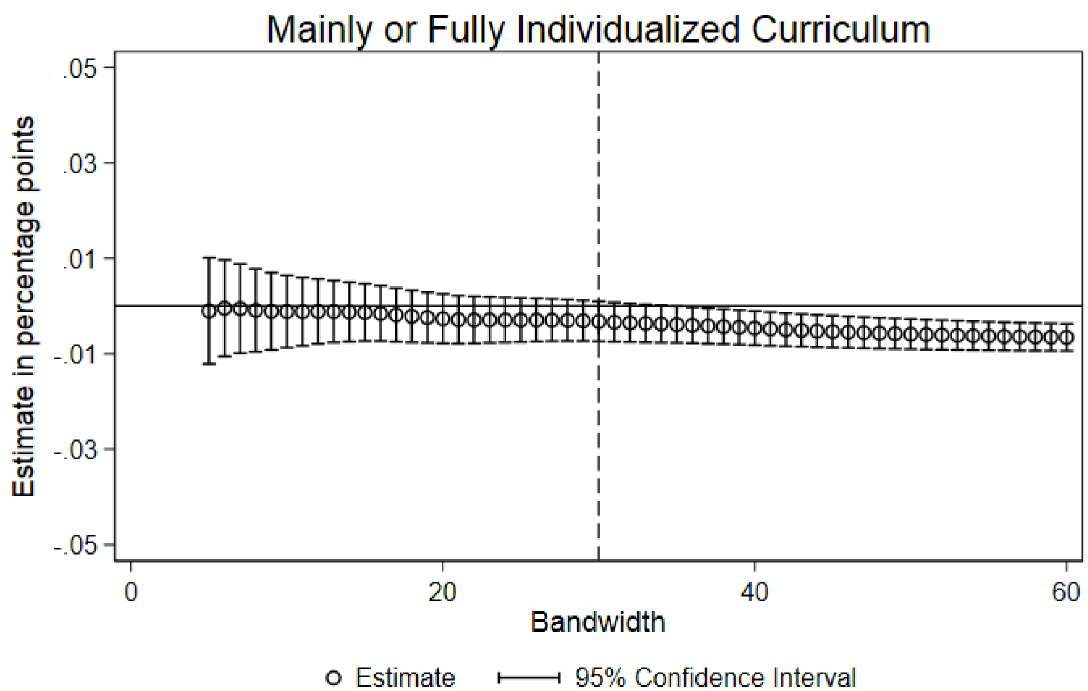

Figure A.2: *Estimation points and 95% confidence intervals for different bandwidths for mainly or fully individualized curriculum. The horizontal axis is the length of bandwidth on one side of the cutoff. Author's rendering of data from Official Statistics of Finland.*

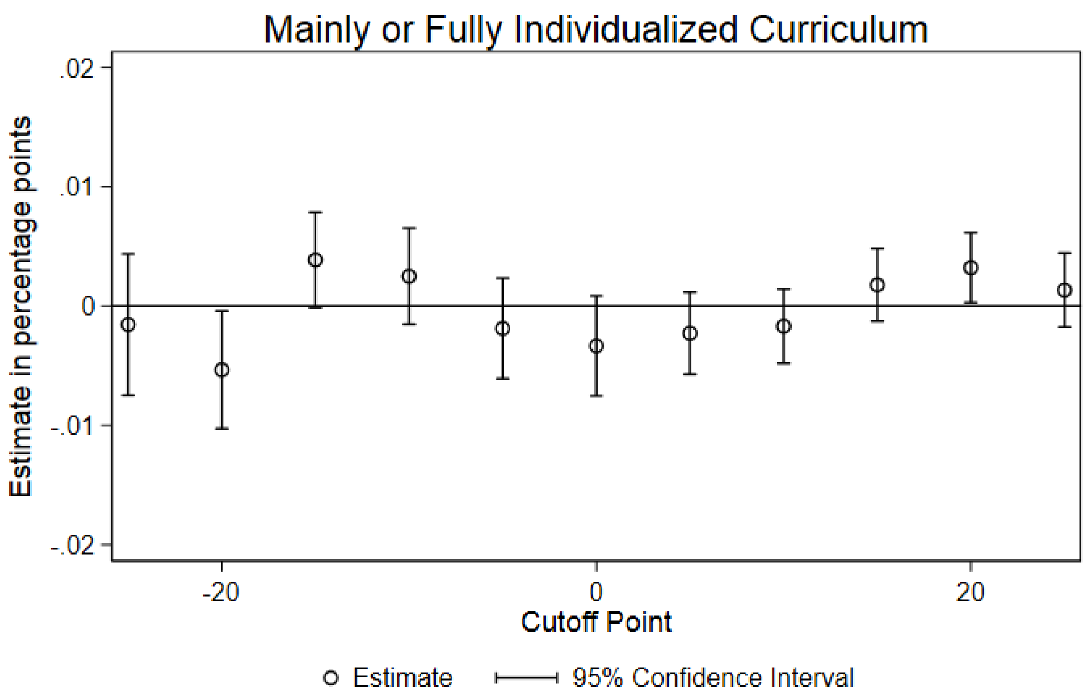

Figure A.3: *Estimation points and 95% confidence intervals for placebo cutoffs for mainly or fully individualized curriculum. The horizontal axis shows the cutoff date used in the estimation. Author's rendering of data from Official Statistics of Finland.*