Unlocking Potential: The Impact of Dyslexia Diagnosis and Support on Educational Outcomes and Well-being

By Søren Albeck Nielsen & Kristoffer Ibsen*

This paper examines how the diagnosis and targeted support of dyslexic children in compulsory schooling impact their educational outcomes and overall well-being. Although dyslexia affects roughly 7% of the population, almost no evidence exists on the differential outcomes of receiving a dyslexia diagnosis compared to remaining undiagnosed. Using Danish data and a regression discontinuity design, our findings highlight the positive impact of a dyslexia diagnosis on short-term, medium-term, and long-term academic performance. We find notable improvements in areas where dyslexic students typically struggle, such as spelling and reading. While proficiency in English as a second language does not show similar gains, there are spillover effects to other subjects, particularly in math and, to a lesser extent, in science. In the long term, dyslexic students demonstrate an increased likelihood of pursuing education after compulsory schooling. An interesting observation post-diagnosis is the initial dip in well-being metrics, which, however, revert to their baseline after a span of two years. Notably, our findings indicate an elevated academic self-perception. These findings emphasize the importance of early detection and support for dyslexic students to help them succeed in the educational system. They also underscore the need for standardized testing and a less deterministic approach to supporting children on the margin of diagnosis.
JEL: H52, I21
Keywords: Dyslexia, reading, academic performance, well-being

> "To the dyslexic student learning to read, books aren't a portal into another world, but a door that keeps slamming in their face."
> Sirin Kale, *The Guardian* (2020)

The fundamental role of reading proficiency in human capital acquisition is indisputable, yet a significant number of children in developed nations struggle in achieving this basic skill. According to the 2019 PISA test, 23 percent of children from OECD countries par-

* TrygFonden's Centre for Child Research & Department of Economics and Business Economics, Aarhus University. We gratefully acknowledge helpful comments and discussions from Eric Bettinger, Elena Mattana, Michael Rosholm, Alexander Koch as well as seminar participants at the Stanford Causal Science Center Conference, VIVE, 2nd and 4th meeting on Quantitative Educational Research in Denmark, and the CESifo / ifo Junior Workshop on the Economics of Education 2024

ticipating in the assessment fell below the minimum reading proficiency expected for their age. Even among the highest-performing countries (excluding four provinces in China), over 10 percent of students failed to reach this crucial benchmark. In the United States, the National Association of Education Progress[1] scores reflect a similarly discouraging assessment of reading achievement. 34 percent of fourth-graders and 27 percent of eighth-graders were in 2019 not reading at a basic level.[2] Understanding why children are unable to attain proficient reading levels is of crucial importance for education policy and advancing our understanding of human capital accumulation.

One plausible reason for these alarming numbers on reading proficiency may lie in our insufficient ability to identify children with reading challenges and the subsequent shortfall in providing effective support systems to address them. This paper focuses on dyslexia, the most commonly recognized and predominant reading disability. Specifically, dyslexia is a genetic neurological disorder that impairs individuals' reading, spelling, and writing abilities. It is estimated that dyslexia affects roughly 7 percent of all children, corresponding to 1-2 children in a typical school class, although this figure is contingent on diagnostic benchmarks (Peterson and Pennington, 2012). Without timely diagnosis and support, these children may confront escalating challenges, experiencing not only academic stagnation and decreasing well-being but also barriers to attaining their full academic potential.

In this paper, we examine how the diagnosis and targeted support of dyslexic children impact their educational outcomes and overall well-being. Despite dyslexia's notable prevalence and significance, there is a remarkable scarcity in evidence regarding the outcomes of receiving a dyslexia diagnosis compared to not being diagnosed. Much of the existing literature primarily revolves around descriptive studies or small-scale interventions that evaluate specific aids with less emphasis on the broader implications of receiving a dyslexia diagnosis (Toffalini et al., 2021). One challenge arises due to the non-random nature of how children are diagnosed with dyslexia. In numerous educational environments, dyslexia diagnoses depend on subjective evaluations that, despite being administered by professionals, are based on referrals from teachers and self-reporting. This inconsistency can lead to children being repeatedly tested throughout the educational system using different criteria, which may, in turn, yield different results (Erbeli, Rice and Paracchini, 2022). Consequently, accurately estimating the impact of dyslexia diagnosis often encounters significant challenges due to the presence of selection biases.

To overcome these challenges, we exploit that Denmark in 2015 implemented a nationwide standardized testing scheme for children at risk of dyslexia in compulsory schooling.

[1] "NAEP Report Card: 2019 NAEP Reading Assessment", National Assessment of Educational Progress (NAEP), *https* : *//www.nationsreportcard.gov/highlights/reading/*2019/ (accessed February 1, 2023).

[2] These numbers had in 2022 increased to 37 percent and 30 percent respectively, indicating that COVID-19 potentially had a negative effect on children's reading skills on average.

The test is individual, computer-based, and conducted under the guidance of a trained instructor with a teaching background. This implementation eliminates the reliance on subjective evaluations and ensures a uniform method of diagnosis across the student population. Leveraging the standardized test, we adopt a regression discontinuity (RD) design to identify the effects of diagnosis and targeted support. Importantly, children on either side of the dyslexia threshold, those barely diagnosed and those narrowly undiagnosed, should exhibit similar characteristics across all dimensions, allowing us to use a RD design to estimate effects for children on the margin of diagnosis. Consequently, whether a child scores just above or just below the threshold on a given test day could arguably be viewed as essentially random.

We estimate a sharp RD to examine the intent-to-treat (ITT) effects of a dyslexia diagnosis in accordance with the standardized testing system in Denmark. Since some children might be diagnosed via non-official tests, our true measure of support comes from their enrollment in a government-provided membership database for dyslexic materials. For enrollment in this database, individuals are required to present proof of their dyslexia. If not diagnosed through the standardized test, this evidence must originate from an alternative testing source. Hence, we use this as instruments for receiving a dyslexia diagnosis in a fuzzy RD design.

Our first empirical finding is that the effect of receiving a dyslexia diagnosis on educational outcomes is positive: we find significant positive effects of dyslexia diagnosis on academic performance, specifically for reading test scores, in the short term. These benefits extend into the medium term, particularly in areas where dyslexic students commonly face challenges, such as spelling and writing. We find spillover effects to math and to some extent science, indicating broader academic improvements. Despite observing positive effects on academic performance, particularly in reading, spelling, and writing, we find no notable improvements in English proficiency, which may be attributed to its status as a second language. Furthermore, in the long term, dyslexic students demonstrate an increased likelihood of pursuing education after completing compulsory schooling. This finding suggests that with appropriate support, dyslexic students can effectively mitigate the challenges associated with dyslexia, enabling them to pursue further education.

Second, we find suggestive evidence that the diagnosis in the long run does not affect well-being. However, we observe a temporary decline across four distinct well-being measures following the diagnosis. However, these measures recover and return to their initial levels after a period of two years. Notably, we find a positive tendency towards a higher academic self-perception, indicating an improvement in students' confidence in their academic abilities over time.

Third, our heterogeneity analysis provides suggestive evidence that early diagnoses in

terms of age appear to offer pronounced long-term advantages, though results differ across academic subjects. Although girls demonstrate notable improvements in the short to medium term, especially in reading, boys tend to reap more significant benefits over time. While children from low-SES backgrounds show immediate improvements, particularly in reading, long-term benefits are primarily driven by high-SES children.

This paper is, to the best of our knowledge, the first paper providing empirical evidence of the causal effects of receiving a dyslexia diagnosis. Our work offers novel evidence about a large group of children that in some contexts may have been forgotten in terms of reading interventions. Our paper plays a pivotal role in raising awareness about dyslexia, challenging stereotypes, and underscoring the importance of early diagnosis and intervention. Most importantly, it demonstrates a notable impact on educational attainment for a traditionally challenging group, potentially yielding substantial economic returns (Heckman, Humphries and Veramendi, 2018). Failing to support dyslexic children risks wasting their untapped educational potential, which in turn leads to a loss of human capital and decreased productivity in society.

This paper also extends the literature on other neurodevelopmental disorders such as attention-deficit hyperactivity disorder (ADHD) and Attention Deficit Disorder (ADD). Studies have indicated that ADHD is negatively associated with human capital accumulation (Currie and Stabile, 2006) and adverse adult labor market outcomes (Fletcher, 2014). However, a similar challenge with these disorders is their complicated and subjective diagnostic process. Consequently, much of the existing research predominantly leans on sibling comparisons or longitudinal analyses.

Our findings underscore the significance of effectively testing, diagnosing, and supporting children with reading challenges. Specifically, they indicate that children who score just above the dyslexia threshold could also stand to gain from targeted support, even if just for a limited duration. Such proactive measures might prevent these borderline cases from falling behind their peers. This study further emphasizes the value of adopting standardized dyslexia evaluations. Uniform assessments not only guarantee consistency across different settings but also ensure that once a diagnosis is made, the child receives timely and effective support tailored to their specific needs.

The reminder of the paper unfolds as follows: Section I focuses on the dyslexia aspect, institutional settings, testing procedures, and the support options when receiving a dyslexia diagnosis. In Section II, we present the data used and any sample restrictions applied. The empirical strategy is detailed in Section III, followed by the presentation of our findings in Section IV. In Section V, we discuss suggestive mechanisms. Finally, Section VI concludes the paper by summarizing the key findings and their implications.

I. Dyslexia, Institutional Settings, and Testing

A. Dyslexia

Dyslexia is a complex condition that is often surrounded by many myths and misconceptions. However, it is a widely recognized fact that dyslexia is a neurodevelopmental disorder that affects the brain's ability to process language, leading to challenges in reading, writing, and spelling (Peterson and Pennington, 2012). Research employing brain imaging techniques finds that dyslexia correlates with diminished activity and dysfunction in vital areas of the brain's left hemisphere (Shaywitz and Shaywitz, 2020). As these areas are pivotal for language processing, their compromised activity clarifies the reading difficulties encountered by individuals with dyslexia. Additionally, other research has highlighted the genetic aspects of dyslexia. Specifically, dyslexia exhibits strong heritability with specific genetic markers linked to the disorder (Doust et al., 2022).

The disorder results in a deficit in phonological processing, which is the ability to discern, recall, and manipulate distinct sounds, also known as phonemes. For instance, distinguishing "bat" from "cat" based on the 'b' sound illustrates phonological processing. This deficit hinders dyslexics from decoding the relationship between letters and their corresponding sounds efficiently (Démonet, Taylor and Chaix, 2004). While non-dyslexic children's decoding between letters and sounds gradually become more efficient and automated over time, and eventually become effortless, dyslexic children may struggle with even the most basic words that they have encountered multiple times (Snowling, Hulme and Nation, 2020). This struggle can make reading a slow and exhaustive process, leading to frustration and feelings of inadequacy. Some dyslexic children may even resort to disruptive behavior or clowning around in class as a way to hide their difficulties. They may even prefer being sent to the principal's office rather than being asked to read aloud in class. Others may feign disinterest in their grades or view school as pointless, failing to realize their full academic potential.

While brain imaging and genetics provide promising insights, they are still infeasible tools for diagnosing dyslexia due to methodological challenges and cost-effectiveness. However, brain imaging and genetics have taken dyslexia from a hidden disability to one that is potentially visible. Still, the lack of clear-cut and objective diagnostic criteria remains the primary obstacle in diagnosing dyslexia (Gibbs and Elliott, 2020). Adding to the complexity is the fact that dyslexia exists on a continuum, rather than as a simple binary condition, making it challenging to establish unambiguous criteria for diagnosis based solely on reading ability (Gabrieli, 2009). Historically, dyslexia has been defined as an unexpected, persistent, and specific learning disability that results in slow and/or inaccurate word recognition and spelling despite adequate instruction and intelligence (Lyon, Shay-

witz and Shaywitz, 2003; Démonet, Taylor and Chaix, 2004). This unexpectedness stems from the typical alignment of reading skills with IQ. In children with dyslexia, however, their reading performance often lags behind their IQ predictions (Shaywitz and Shaywitz, 2020). Although earlier diagnostic approaches prioritized this discrepancy, it is essential to recognize that individuals with dyslexia, regardless of IQ, encounter the same core phonological challenges (Snowling, Hulme and Nation, 2020).

The diagnosis of dyslexia is further complicated by the fact that a phonological deficit is not the only impairment that may be observed. Dyslexia is frequently accompanied by other psychiatric disorders like ADHD, ADD, or anxiety. For diagnoses, it is, therefore, crucial to distinguish dyslexia from other learning disabilities that can impact reading.

The lack of consensus on its definition and objective diagnostic measures contributes to uncertainty around its prevalence (Démonet, Taylor and Chaix, 2004). One approach commonly used is to identify individuals who score 1.5 standard deviations below the average reading achievement for their age, which typically yields a dyslexia prevalence of around 7 percent (Peterson and Pennington, 2012). This approach involves the establishment of a cut point on a continuous distribution, and it is evident that the placement of this cut point significantly impacts prevalence estimates (Wagner et al., 2020). In line with this, other studies report a broader range of prevalence estimates, from 5 to 17.5 percent, depending on how dyslexia is defined and whether other impairments are considered (Shaywitz, 1998). Based on this, there is no definitive prevalence estimate that can be considered universally correct. Instead, the prevalence is influenced by the severity of the disorder.

While dyslexia cannot be cured, many dyslexics can mitigate their struggles through extra effort and the use of alternative learning strategies. By putting a greater dependence on right-hemisphere pathways, which plays a crucial role for visualization and memory, individuals with dyslexia can effectively compensate for their difficulties and potentially reach a level where their condition is not noticeable to others (Hoeft et al., 2011). The process of visualization, a right-hemisphere function, aids in the creation of mental images. Dyslexic individuals can use this process to visualize words or concepts and thereby enhance their understanding and recall. Memory pathways, another right-hemisphere function, can assist in memorizing specific word forms and spelling patterns. An early implementation of these coping strategies can make a significant difference. As such, early detection of dyslexia becomes a crucial step in understanding and addressing children's learning challenges, preventing dyslexia from becoming a lifelong challenge for these individuals (Gabrieli, 2009).

B. Institutional Setting

In this paper, we focus on children in compulsory school in Denmark. Compulsory schooling begins in August in the calendar year in which the child turns six and lasts 10

years, consisting of a preschool class (also termed 'Grade 0') in the first year followed by Grades 1-9 with an optional Grade 10.[3] Generally, all education from compulsory schooling to higher education is free of charge. It is, however, possible to attend private schools with tuition fees for compulsory schooling. 22 percent of the children who had their first day of school in 2021 started in a private school. Children are typically grouped into classes when entering preschool and usually stay in the same class until the end of Grade 9, unless they change schools.[4] National laws dictate that the maximum class size is 28 students with an average of 21 in 2021, which is similar to the OECD average. While the preschool class has a grade teacher, Grades 1-9 are taught by subject-specific teachers who often stay with the same group of children for several years.

Compulsory schooling ends with mandatory final exams in Grade 9 on seven subjects.[5] Written examinations are set and marked centrally, while centrally appointed examiners assess the oral examinations together with the subject teacher. In addition to the final exams, Danish schools are required to conduct a number of standardized tests and well-being surveys for different grade levels. The tests are adaptive and computer-based, meaning that they constantly adjust the next question to the level of the individual child depending on right or wrong answers. The survey is also conducted electronically with one designated teacher responsible for conducting the survey in the class. The frequent testing allows teachers to track the learning progress and well-being of each child and identify students who may need further assessment for special needs such as dyslexia. Appendix A shows the frequency of the tests and survey.[6] After compulsory schooling, the child can enroll in either high school or vocational school. While a high school degree provides access to higher education, a vocational degree primarily qualifies for skilled jobs in the labor market.[7]

The Danish school system is built on the principle of inclusion, meaning that education is made available to all individuals, regardless of any difficulties they may have. This principle entails that the educational support in terms of special education is tailored to each child based on a thorough assessment of the individual needs. Specifically, whether a child's development requires special education is based on an assessment from a public educational-psychological counselling service that also evaluates the nature of the needs.

[3] 51 percent of the children attended Grade 10 in 2021. Of these children, 65 percent attended a boarding school, which is common for 10th graders in Denmark.

[4] A few schools reshuffle classes in Grade 3 and/or Grade 7.

[5] Five exams are mandatory, while two exams are randomly drawn from a pool of subjects. Mandatory exams are: written and oral exams in Danish, a written exam in math, an oral exam in English, and an oral common exam in physics/chemistry, biology, and geography. Note that the written exams in Danish and math consists of subexams.

[6] We refer to Beuchert and Nandrup (2018) for specific details about the content and testing procedures, while Andersen et al. (2020*b*) describe the details of the well-being survey.

[7] A vocational degree can also with certain requirements provide access to specific higher education degrees that take practical approaches.

Ultimately, the school principal in consultation with the child and parents decides whether a child should be referred to special education. If referred to special education, the specific educational support is reassessed at least once a year. 3.56 percent of all children with dyslexia received special education between 2015 and 2019, but schools also assign special education to address issues such as ADHD, ADD, intellectual disabilities, and hearing impairments. Students in need of more than 9 hours of special education per week will typically receive the extra support outside the classroom since the child may receive special education to supplement or replace participation in regular education in one or more subjects.[8] Alternatively, all teaching for the child may be provided in a separate class, either at the same school or at certain schools that specialize in providing education for students with special needs.

C. Testing for Dyslexia

In 2015, the Ministry of Children and Education in Denmark implemented a free of charge nationwide dyslexia test which provides a standardized method to identify dyslexia throughout the educational system in Denmark. The standardized test ensures that the same procedure is used to identify dyslexia and does not depend on the assessor. Non-standardized testing of dyslexia occurs in many countries, often resulting in repeated testing of the children throughout the educational system with often different criteria being used (Erbeli, Rice and Paracchini, 2022). Another advantage of the Danish dyslexia test is that it is designed such that it can be adjusted to all grade levels.

As a general rule, children can only take the test if teachers at the school have reasonable suspicion that the child is dyslexic. This suspicion typically comes from other academic tests or from observations in the classroom. Parents have to consent before the test can be conducted. From 2017, parents could demand that their children take the test at least one time during compulsory schooling if their child attended a public school. The one time restriction follows from the worry that taking the same test may result in an inflated test score. However, retesting may indeed be necessary. A student who did not show any signs of dyslexia in early schooling can turn out to have the disorder years later. Early well-conducted teaching that fits the student's needs can hide the disorder, but when reading and writing requirements over time increase to a level that supersedes the teaching level, or simply if the teacher changes, it could show up. In other words, the disorder does not go away, but the extent of it can be reduced with the right help.

The primary focus of the Danish dyslexia test is to identify difficulties related to phono-

[8]Students in need of less than 9 hours of of special education per week usually follow the same schedule as their classmates and receive the extra support in the classroom. This extra support typically consists of differentiated teaching or by assigning a teacher's aide to the class room that can help the individual child and the class in general. A teacher's aide can in particular be helpful in classrooms with behavioral problems (Andersen et al., 2020*a*).

logical processing. The test usually takes around half an hour and consists of a survey as well as three subtests. While the two first subtests test the presence of a phonological deficit, the last one tests the validity of the two previous subtests. We describe the exact components of the subtests in Appendix B. The test is computer-based, and test instructions are read out loud by the computer. The dyslexia score is based on a mean of the first two subscores. Each of the subscores is adjusted for the number of correct answers and the time spent on each exercise (see Appendix B for the specifics of the test scores). This correction means that you can have a high percentage of correct answers, but if you only try a few exercises within the time limit, you will get a low score. One of the characteristics of dyslexics is that they have slow decoding skills, so even though they get the exercises right, they may be significantly slower than they are supposed to, which is something that the test needs to take into account.

If you receive a score below a certain age-dependent threshold, you are diagnosed as dyslexic. If you score above, you are either assessed as having an inaccurate phonological coding or as not dyslexic. Being labelled as having inaccurate phonological coding suggests that you should get some extra attention, but it is uncertain whether and to what extent this actually happens. Based on the test, approximately 8 percent of the population is expected to get the diagnosis of being dyslexic. Another 7 percent is expected to have inaccurate phonological coding. While dyslexia is a specific diagnosis, it is extremely difficult to diagnose precisely. Given that the diagnosis is based on a test that takes around half an hour, it does indeed seem random whether you on a given test day score just below or above the cutoff of being diagnosed with dyslexia.

Trained test instructors with a background in teaching assist in carrying out the test. The main purpose of the presence of the test instructor during the test is to ensure that it is carried out in a reliable manner. For example, the reliability may be doubtful if the student attempts many questions, but only answers a few of them correctly. This behavior could indicate that the student just guessed randomly on the questions without considering the options in order to get it done quickly. To disregard this concern, the test instructor observes the whole test. After the test, the test instructor evaluates its validity based on the student's behavior during the test and the results of the third subtest. If validated, the instructor can directly access the dyslexia score from an online portal.[9] If the test is not approved, the student can retake it when deemed appropriate; however, this rarely occurs.

The dyslexia test, while valuable, has its limitations. First, its reliability is based mostly on data collected between March and June, as it is believed these months best represent grade-level reading skills due to their position at the end of the academic year. However,

[9]We only include results from validated tests in our data.

this means tests taken outside this period might be less reliable. Second, the test has not been fully vetted for three specific groups: children receiving special education in separate classrooms, children with primary issues other than reading (such as ADHD or cognitive impairments), and children with limited proficiency in Danish as a non-native language. These limitations underscore the necessity for careful sample selection as described in section II.

It is worth noting that alternative dyslexia tests exist apart from the Danish dyslexia test. As the Danish dyslexia test is not supposed to be retaken, some children, parents, or teachers who are not willing to accept the test result may opt for another dyslexia test. These tests are informally acknowledged by schools as indications of dyslexia, and consequently, children who receive other results from these tests may receive similar assistance to those diagnosed through the Danish dyslexia test. Unfortunately, we are not able to observe whether these additional assessments are undertaken. However, we are able to observe whether the children enroll in a government-provided membership database for dyslexic materials. Enrollment in this database serves as the most reliable proxy available to determine access to dyslexia support. Importantly, upon enrollment, individuals must provide evidence of their dyslexia, which then has to come from a test other than the Danish dyslexia test if the evidence is not obtained through that particular test. Based on anecdotal evidence, it is suggested that being a member of the database is the most significant and helpful resource for individuals with dyslexia. Consequently, we anticipate that all children who are identified as dyslexic, no matter the exact way, would become members of the database.

D. Receiving a Dyslexia Diagnosis

Upon being diagnosed with dyslexia, students benefit from a wide array of support options tailored to their unique needs such as digital tools, individualized teaching adaptations, and modified exam conditions. While our data set identifies who receives a diagnosis, it does not delve into the specifics of how these support options are utilized.

In the early stages post-diagnosis, the test instructors play a pivotal role in identifying specific challenges and facilitating access to digital tools. Among these tools, word processing programs with features like spellcheckers and auto-correct are quite common. However, text-to-speech tools stand out as particularly important, especially for students leaning heavily on visualization and memory to mitigate the disorder. This combined auditory and visual approach significantly enhances their comprehension by enabling them to listen to the content while simultaneously following along visually. One of the primary purposes of the government-provided database for dyslexic materials is indeed to record all school books so they can be read aloud to dyslexic children.

In the classroom, the dyslexia diagnosis can also lead to a shift in how teachers and peers perceive the child. When they are aware of a child's dyslexia, they may be more patient and accommodating, which can help reduce the child's anxiety in the classroom and improve their self-esteem. Teachers can modify their teaching strategies and provide accommodations, such as providing reading materials in alternative formats, to help the child learn more effectively. However, if the teachers are unfamiliar with the disorder, their well-intended efforts may fall short of providing the necessary support, or worse, unintentionally reinforce myths and biases about dyslexia.

Beyond the classroom, special conditions may also be present for dyslexic students at the final exams. First, dyslexic students may use specific digital tools at the final exams, even though they are not allowed for all other students. However, the digital tools naturally extend what dyslexic students use everyday. Second, dyslexic students may be granted additional time to complete the exams due to the increased time it typically takes them to complete tasks using the digital tools. Thirdly, because of the extended time provision, students with dyslexia might be placed in a separate exam room to avoid disruptions from other students leaving when the standard time has elapsed. The school principle ultimately decides on the specific conditions for each student based on an individual assessment of the student and input from parents and teachers. Under special circumstances, the school principal can also exempt a student from taking one or more of the final exams if the exam is found to be inappropriate due to severe disabilities.[10]

In terms of well-being, receiving a dyslexia diagnosis can be a turning point in a child's life. Having a name for their struggles and an explanation for their difficulties can bring immense relief to dyslexic children. Knowing that their difficulties are due to a neurological condition and not a lack of effort or intelligence can help them feel less discouraged and more hopeful about their future academic prospects. On the other hand, while a dyslexia diagnosis can be empowering, it can also be a source of concern and worry. Some children, parents, and teachers may view dyslexia as a permanent, unchangeable condition, leading to a belief that progress is not possible (Gibbs and Elliott, 2020). Moreover, there could be some stigma surrounding dyslexia, which may cause some individuals to feel embarrassed or ashamed of their difficulties, which could result in a reluctance to seek help or disclose their condition to others.

While we have insights into students who receive a diagnosis based on the standardized cutoff, it remains uncertain whether those marginally above the threshold are also benefiting from the support options. The nuances in how and if these near-threshold students access and use available resources are not explicitly captured in our data either. If those

[10]This exemption may have consequences for enrolling in high school since a key requirement is that the student have attempted all exams in Grade 9. The student can, however, still enroll in high school based on an entrance exam and an interview.

just to the right of the cutoff (those who barely miss out on the diagnosis) are still able to access and benefit from the support options, then the estimated effects we measure would be diminished. This would make our estimates more conservative, providing a lower bound for the effects of the diagnosis and the subsequent support.

II. Data and Sample Restrictions

Our analysis combines several data sources that are interconnected through an anonymous social security number assigned to each individual. The data on dyslexia is from two distinct and previously unavailable sources. First, we have obtained data from the Ministry of Children and Education on the Danish dyslexia test that contains all validated tests between 2015 and 2019. This data set includes the dyslexia scores and corresponding test dates. Interestingly, it also contains the answers to the survey that the tested children take before the actual test. From the nine included questions, we can, for example, infer whether they already receive additional support in school, and whether they themselves think that they are dyslexic. Second, we have gained access to the government-provided membership database for dyslexic materials.[11] This database enables us to observe the enrollment of children with dyslexia. To enhance our analysis, we link these unique data sources with administrative registers provided by Statistics Denmark. Among other things, the registers contain information on socioeconomic status and school-specific characteristics, and it also allows us to match the children with their parents, siblings, and classmates. We are able to observe the registers until the end of 2021.

Importantly, the administrative registers also contain data on the mandatory standardized test administered in Grades 2-8 (available since 2010) and results from the final exams at the end of Grade 9. We also observe the national well-being survey (available since 2015), which contains responses to a survey measuring well-being for grade 4 to 9 children in public schools. Additionally, we gather information on the subsequent paths pursued by students after completing compulsory education.

Leveraging this data, we construct our primary outcome variables to assess the educational achievements of each child in what we classify as the short run, medium run, and long run. Initially, we examine the short-term impact by evaluating performance on the first reading and math tests subsequent to the diagnosis. Second, we analyze the medium-term effect by assessing performance on the final exams in Grade 9. Lastly, we examine the long-term outcomes by investigating whether children continue within the educational system two years after completion of compulsory schooling.

As secondary outcomes, we examine the impact on various well-being measures. Specifically, we focus on academic self-perception, feelings of support, academic well-being, and

[11] Known as NOTA in Danish.

social well-being[12]. Except for academic self-perception[13], these measures consist of a combination of 40 questions that have been validated through a Principal Component Analysis (PCA) conducted by the Ministry of Children and Education. The feelings of support encompass the child's perception of teachers' assistance and support as well as their own motivation and participation. Academic well-being includes questions related to academic skills, concentration, and problem-solving abilities, while social well-being focuses on the sense of belonging to the school and class, feelings of safety, and experiences of bullying. Furthermore, we assess any changes in the frequency of absences.

For all continuous outcomes, we standardize within grade, year, and each subject or survey-item to mean zero and a standard deviation of one. For aggregate measures (i.e. the reading test score, GPA, and well-being outcomes), we then calculate the average and standardize it within the grade and year to mean zero and a standard deviation of one. The latter standardization allows us to readily interpret coefficients in standard deviation units and, thus, assess whether the results are comparable to the effect sizes of other studies.

Appendix Table C1 provides an overview of the sample selection process and the resulting sample size. The initial pool consisted of 67,533 children who took the national dyslexia test between 2015 and 2019. To refine the sample, several criteria were applied.

First, children who did not receive regular classroom teaching, including those in extensive special education or not enrolled in private or public schooling, were excluded. Secondly, only participants tested during the validated test time (spring) were included. Additionally, children who were already members of the government-provided dyslexia database prior to the national dyslexia test were excluded. Children recently diagnosed with a mental disorder were also excluded, as the test is not validated for this group. Finally, children with at least one primary outcome measure were retained, resulting in a final sample size of 31,610 children. This rigorous sample selection ensures that our findings are based on a well-defined and valid sample of children with dyslexia.

III. Empirical Strategy

Our goal is to estimate to what extent diagnosing and providing targeted support to children with dyslexia impacts their academic performance and overall well-being. To accomplish this, we exploit the diagnostic capabilities of the Danish dyslexia test. This specific test equips us with the essential scores required to draw a comparison between dyslexic and non-dyslexic children. However, the severity of dyslexia can vary to a large

[12]Before the survey, the teacher emphasizes that the primary objective of the survey is to enhance the overall well-being of everyone within the school community. Importantly, the teacher also assures the students that their responses will be kept strictly confidential and will not be shared with their parents, teachers, or anyone else within the school. At the same time, it is also stated that they should respond honestly and that they can have the questions read out loud if they have reading problems or be helped in other ways.

[13]Consists of the questions: "I am doing well in school work" and "I am making good academic progress in school"

extent, leading to substantial differences between children on opposite ends of the continuum.

We use the dyslexia cutoffs in a regression discontinuity (RD) design, which enables us to compare children who fall just above and below the cutoff for dyslexia diagnosis. The fundamental assumption is that children immediately above and below the cutoff are similar, except for their assigned treatment being the diagnosis. If children are unable to precisely manipulate their treatment status around this cutoff, it can be argued that the variation in treatment near the cutoff is as good as random and hence fits the essence of a randomized experiment (Lee and Lemieux, 2010). For each outcome y_{it} on academic performance and well-being, we start by estimating the following model:

$$y_{it} = f(score_i) + \beta \cdot \mathbb{1}(score_i < 0) + \delta X_{it} + \epsilon_{it} \tag{1}$$

where y_{it} denotes the outcome for child i at time t, $score_i$ is the running variable and is the test score relative to the age-specific dyslexia cutoff of child i, $f(\cdot)$ is a first-degree polynomial function in distance to diagnosis cutoff that is permitted to vary on both sides of the cutoff, $1(score_i < 0)$ is an indicator variable for child i being diagnosed with dyslexia in the dyslexia test, and X_{it} represents a vector containing pre-determined covariates. Under the maintained assumption that there is no discontinuity in the distribution of ϵ_{it} at the cutoffs, the parameter of interest, β, represents an intention-to-treat (ITT) estimate in a sharp RD design of the effect of being dyslexic in the standardized test and receiving the subsequent support.

Our main results are based on local linear regressions in which we employ a triangular kernel that assigns less weight to observations that are further away from the threshold. The standard errors are clustered on the dyslexia test score (Lee and Card, 2008). As a robustness check of our estimates, we also report the bias-corrected estimates with robust confidence intervals (Calonico, Cattaneo and Titiunik, 2014). In this scenario, we calculate standard errors from a specification that in our case incorporates a second-degree polynomial for the running variable. We compute optimal bandwidths based on Calonico, Cattaneo and Titiunik (2014). The optimal bandwidth falls between 4.34 and 7.62 with a mean of 5.34 for the main outcomes that we study (see Appendix Table D1). Thus, we use a bandwidth of 6 points on either of the dyslexia cutoff to ensure that children on either side of the dyslexia cutoff are similar. In the following sections, we show that our results are robust to alternative functional forms, specifications, bandwidths, and whether or not covariates are included.

While the reduced-form coefficients (i.e. the sharp RD design) are interesting in their own right, there is a concern regarding potential always-takers, where children might acquire the dyslexia diagnosis through alternative testing methods. Consequently, this situation

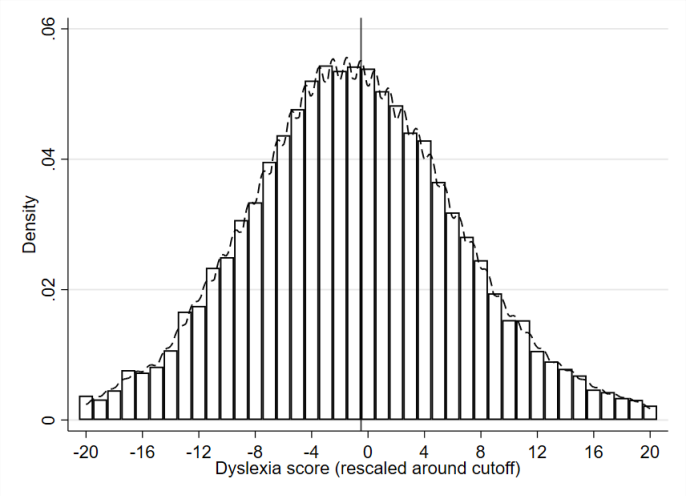


Figure 1. : Frequency of observations around the dyslexia cutoff

could lead to an underestimation of the impact of being dyslexic and receiving support. To address this issue, we incorporate a fuzzy RD design as a means of accounting for the non-compliance. This approach can be thought of as an instrumental variable strategy where we might be able to interpret our estimates as a local average treatment effect (LATE). However, to be able to do so, we additionally need to satisfy the assumption of monotonicity and the exclusion restriction (Imbens and Angrist, 1994). These results should therefore be interpreted cautiously. The monotonicity assumption requires that crossing the dyslexia threshold does not make an individual less likely to being provided dyslexia support. This assumption of no defiers seems likely to hold in our setting. Additionally, the exclusion restriction requires that crossing the dyslexia cutoff affects outcomes only through the diagnosis and provided support. In contrast, an advantage of the reduced-form model is that it requires fewer assumptions. In the absence of any manipulation of the assignment variable and in the absence of other discontinuous changes occurring at the cutoff score (which we check in the next section), the reduced form consistently captures the effect of being diagnosed with dyslexia versus not being diagnosed.

The relevant first stage for this fuzzy RD captures the jumps induced by the dyslexia cutoffs in terms of the take-up of provided support. When estimating the first stage, we use the same parametric functional form imposed in the reduced form.

A. Validity of the RD Design

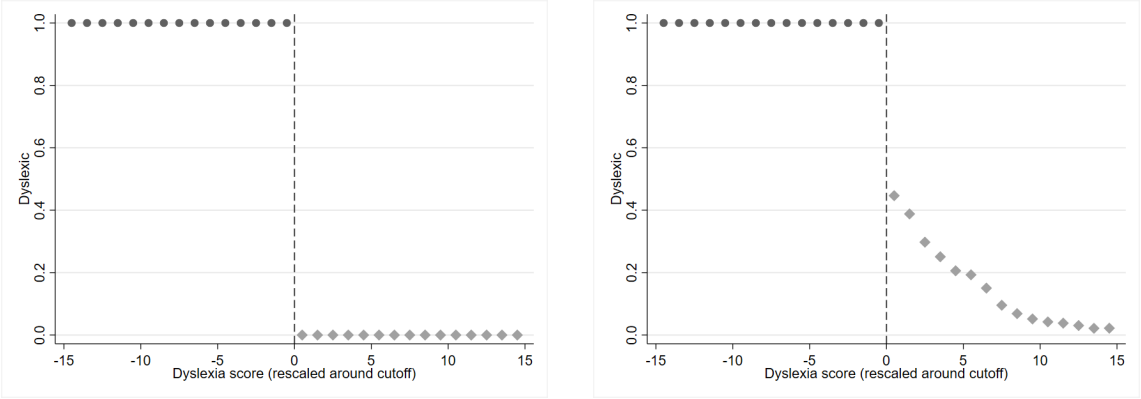
A fundamental concern in any RD design is whether individuals are able to manipulate the assignment variable (Lee and Lemieux, 2010). In this setting, this assumption would

be violated if students, schools, or test examiners could manipulate the test score near the threshold. As the test score is based on an algorithm, and the dyslexia threshold is age-specific, it is extremely difficult to know the exact number of questions the student must answer correctly to pass or fail the test within a specific time frame. Thus, this scenario seems unlikely on a test attempt. We present two types of evidence to support this statement.

First, Figure 1 shows the frequency of diagnoses around the cutoff. The shape of the frequencies is to a high extent a result of our sample selection. On the left side, children with the lowest test scores on the dyslexia test are for the most part already enrolled in substantial special education and thus not included in our sample. On the other side, children are not supposed to take the test if the teachers do not have suspicion that they in fact are dyslectic, which is why we do not observe many children far to the right. In the case of manipulation, we would observe a discontinuity in the density at the cutoff. To provide evidence that there is no manipulation of the running variable at the cutoffs, we conduct a density test based on McCrary (2008) and do not find any evidence of a discontinuity in the frequency of test scores at the dyslexia cutoff. Specifically, we find that the frequency discontinuity at the cutoff is -0.22 with a p-value of 0.83. This evidence provides reassurance that the scores most likely were not manipulated in our context.

Second, we test whether predetermined characteristics are smooth around the cutoff. Appendix Table D2 reports the results of estimating model 1 using baseline characteristics of children and their parents as dependent variables.[14] Column 1 provides the conventional point estimate from the local-linear regression, while columns 2 and 3 present the bias-corrected estimate and the robust standard error (Calonico et al., 2019). Finally, column 4 reports the mean of the covariate in the sample for observations to the right of the cutoff. These results suggest that we generally can verify this condition. However, there are very few characteristics with marginally significant discontinuities at the cutoff. Specifically, factors such as residing with both parents, achieving a higher standardized math score, having an employed father, and a father with relatively higher earnings appear to demonstrate a discontinuity at the specified cutoff. These factors seem to reduce the likelihood of being dyslexic. While chance may account for certain discontinuities, it is plausible that the suggested discontinuities can be attributed to over-referring of certain children due to disruptive behavior at least in the case of having separated parents. For example, Ferman and Fontes (2022) find that teachers tend to favor better-behaved students. This preference, among other factors, can help explain a significant portion of grading discrimination against boys. For all of the remaining 33 predetermined covariates, there is no evidence of a discontinuity around the cutoff at a 5% significance level.

[14]When estimating model 1 for balancing tests, we do not include any other covariates.



(a) Sharp RD (b) Fuzzy RD

Figure 2. : RD Design

The figure illustrates the distribution of individuals diagnosed with dyslexia around the cutoff point. On the left-hand side (LHS) of the figure, we present a visual representation of the sharp design, highlighting the effects of the cutoff on the diagnosed population. On the right-hand side (RHS), we depict the fuzzy design, which takes into account the presence of always-takers, i.e. individuals who receive a dyslexia diagnosis through alternative methods.

Overall, the analyses conducted in this section suggest that there is no manipulation of the running variable around the dyslexia cutoff, nor are there any discernible discontinuities in the observable characteristics of children for almost all covariates. In Section IV.D, we further find that our results are robust to including or excluding all covariates in our main specification. Thus, we argue that the variation in the treatment will be as good as random in the neighborhood around the discontinuity cutoff.

IV. Results

This section presents our main empirical findings. We begin by reporting first stage estimates for how taking alternative dyslexia tests translates into getting dyslexia support. We then present results for how dyslexia impacts academic performance and well-being measures before turning to a variety of robustness checks.

A. First Stage

As a reminder, we have a sharp discontinuity for the diagnosis of dyslexia through the Danish dyslexia test where everybody below the cutoff is dyslexic, and everybody above the cutoff is not. This is illustrated in Figure 2a. We use being provided dyslexia support to scale our reduced form estimates using a fuzzy RD.

We begin by documenting the relationship between being dyslexic and obtaining dyslexia

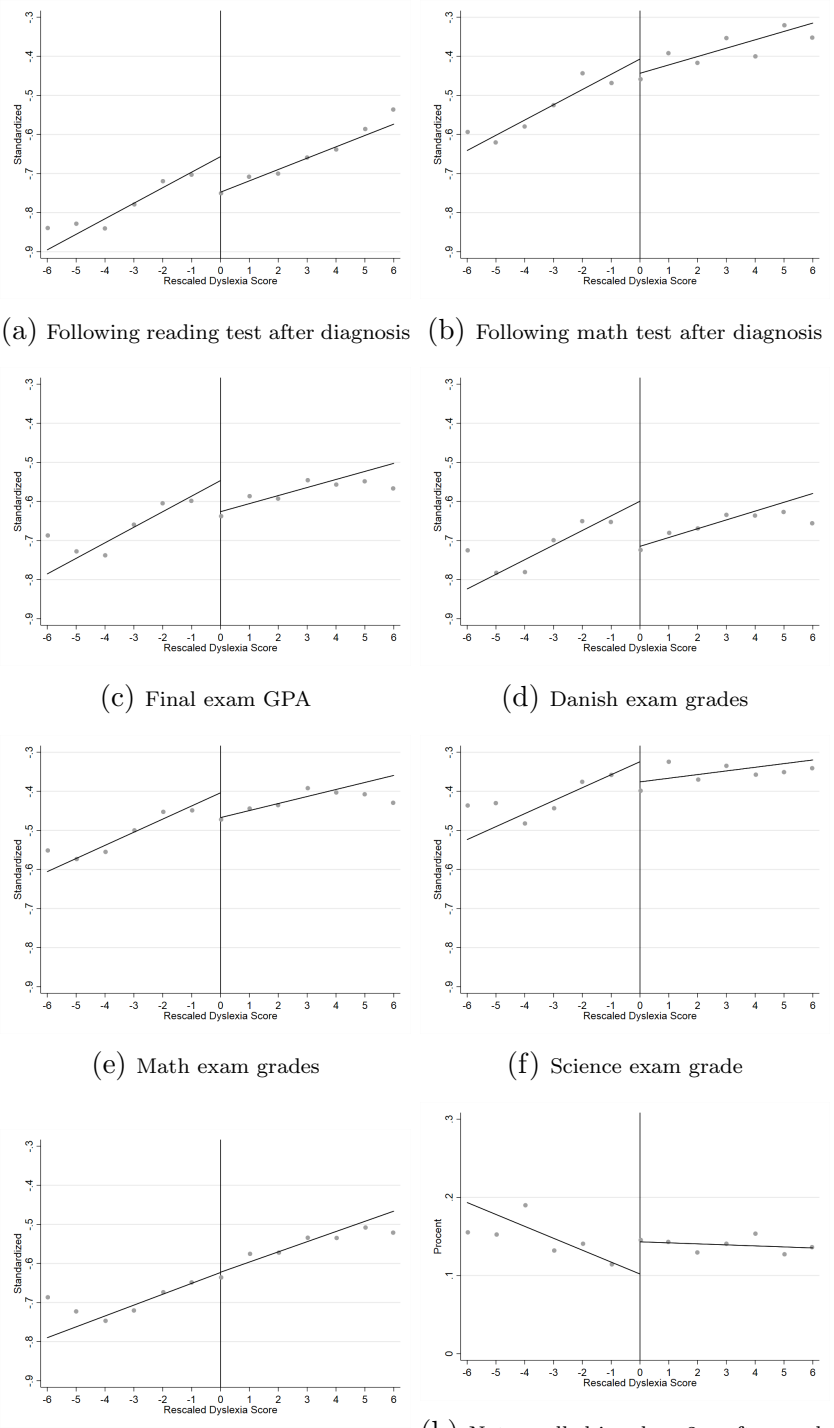
support. Panel (b) of Figure 2 depicts the probability of obtaining dyslexia support. Everyone to the left is diagnosed with dyslexia, while everyone to the right (initially) is not diagnosed with dyslexia. Mechanically, we assume that everyone to the left is getting support, as they due to the diagnosis are automatically assigned support in terms of specialized learning plans, digital tools, and other resources in the educational environment, e.g. getting extra attention from teachers. As detailed in Section I, there are instances where children receive diagnoses through alternative tests. Our most reliable indicator for whether they subsequently receive dyslexia support hinges on their enrollment in the government-provided membership database for dyslexic materials. This explains why children to the right of the dyslexia cutoff may also receive support. There is a descending slope to the right of the cutoff, meaning that children further away from the cutoff are less likely to obtain dyslexia support. This makes sense since children near the cutoff, due to the test's random nature, could be marginally dyslexic if tested on a different day. Put differently, these children might show symptoms of dyslexia in their daily lives, even though they may not be identified as dyslexic based on the test results.

The enrollments into the membership database are, however, not necessarily random, because who chooses to be retested through an alternative test is endogenous. This endogeneity introduces complexities in establishing a causal effect due to unobserved factors that influence the decision to take an alternative dyslexia test. Specifically, parents, who may be aware of genetic risk factors, and teachers, who closely observe the child's performance on a daily basis, are more likely to advocate for support based on information that researchers may not have access to. In light of this information, certain parents and teachers may persistently advocate for additional tests if the Danish dyslexia test only marginally fails to diagnose their children with dyslexia. Hence, it is crucial to instrument for obtaining dyslexia support (which is not random) with being diagnosed with dyslexia (which is quasi-random near the cutoff) to account for selection biases.

Column 1 in Table 1 reports the jumps for each of the educational outcomes. The estimated jumps are sizeable and homogeneous across the the outcomes. The estimates are highly significant, and thus, there is no indication that there will be a weak instrument problem with our fuzzy RD. For all outcomes of interest, the estimate is almost 53 percentage points.

B. Educational outcomes

Figure 3 provides visual evidence on the relationship between the diagnosis of dyslexia and educational outcomes observed for dyslexic children. Each figure shows first-degree polynomial regression lines based on model 1 without covariates. The plotted points are the conditional means of the dependent variable for children in a one-unit bin-width. Panel



(a) Following reading test after diagnosis (b) Following math test after diagnosis

(c) Final exam GPA (d) Danish exam grades

(e) Math exam grades (f) Science exam grade

(g) English exam grade (h) Not enrolled in educ. 2 y after graduation

Figure 3. : Educational Outcomes in the Short, Medium, and Long Run

The figure displays separate first-degree polynomial estimates around both sides of the dyslexia cutoff. Each dot in the plot represents the average outcome for the dyslexia scores closest to the cutoff, providing insight into the impact of the cutoff on the observed outcomes.

(a)-(b) in Figure 3 focus on the short-run outcomes in terms of the next standardized tests in reading and math, panels (c)-(g) show the medium-run outcomes, while panel (f) display whether the child is in education two years after finishing compulsory schooling.

Of particular interest is the assessment of improvements in students' reading, spelling, and writing skills both in the short and medium run. Furthermore, it is interesting to explore the extent to which these improvements extend to other subjects. Enhanced foundational skills are expected to positively impact performance in various courses. Based on Figure 3, there is compelling evidence of a positive influence on reading test scores in the short run in panel (a) as well as in the medium run in the final exams in Danish in panel (d). Furthermore, in panel (c), there are indications of a discontinuity in the effect on GPA. This increase in GPA seems to a large extent to be the result of the positive effect in Danish. However, there appears to be spillover effects from improved Danish skills to both math in panel (e) and science in panel (f), although not to the same extent as observed in Danish. However, there does not seem to be the same spillover effect to math in the short run in panel (b). Similarly, there is no spillover effect observed in English. Moreover, in the long term, there are indications that children are more inclined to persist in the education system, as evident in panel (h).

Column (2) in table 1 presents the corresponding regression results from our baseline model. Each estimate reports the estimated coefficient from a separate local-linear regression with a triangular kernel, bandwidth of 6, and including covariates. We report bias-corrected estimates in brackets and robust standard errors in parentheses. We also report the mean of the outcome for observations above the cutoff and the effective number of observations used in the regressions (referred to as N). Consistent with the graphical evidence, we find statistically significant coefficients for the standardized tests following the diagnosis. Specifically, the reading score is 0.101 standard deviations higher, while the math score is 0.055 standard deviations higher. In the medium run, we further find an ITT effect of being dyslexic on the GPA in the final exams of 0.068 standard deviations. Closer inspection of Table 1 shows that the effect is especially driven by an improvement in Danish of 0.092 standard deviations. As the dyslexia support targets skills in reading, writing, and spelling, this suggests that there is a direct effect of the dyslexia support. In addition, the ITT generates a statistically significant improvement on math and a marginally significant improvement on the exam in science. In contrast, no significant effect was observed in English performance. This is consistent with the availability of dyslexia resources primarily in Danish and the shorter duration that students have had to develop coping strategies, such as memorization, for English. Thus, the limited improvement in English outcomes is contextually understandable.

The Danish and math exams consist of multiple subexams. Appendix Table E1 shows the regression results for each of the subexams. While the other subexams are mandatory, the

Table 1—: Educational Outcomes in the Short, Medium, and Long Run

	First stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean (4)	N (5)
Reading score	**0.522**	**0.101**	**0.191**		
	[0.527]	**[0.105]**	**[0.201]**	-0.665	13,283
	(0.005)	(0.018)	(0.034)		
Math score	**0.522**	**0.055**	**0.104**		
	[0.528]	**[0.062]**	**[0.119]**	-0.390	12,068
	(0.004)	(0.022)	(0.042)		
GPA	**0.527**	**0.068**	**0.129**		
	[0.530]	**[0.081]**	**[0.153]**	-0.580	10,879
	(0.008)	(0.015)	(0.029)		
Danish	**0.527**	**0.092**	**0.174**		
	[0.530]	**[0.099]**	**[0.188]**	-0.664	10,825
	(0.008)	(0.013)	(0.026)		
Math	**0.527**	**0.071**	**0.133**		
	[0.529]	**[0.079]**	**[0.150]**	-0.428	10,810
	(0.008)	(0.011)	(0.022)		
English	**0.525**	0.005	0.009		
	[0.528]	[0.016]	[0.030]	-0.560	10,617
	(0.008)	(0.023)	(0.044)		
Science	**0.525**	**0.040**	**0.075**		
	[0.528]	**[0.063]**	**[0.120]**	-0.355	10,465
	(0.008)	(0.024)	(0.046)		
No educ. 2 years after	**0.527**	**-0.040**	**-0.076**		
	[0.529]	**[-0.047]**	**[-0.090]**	0.140	8,143
	(0.007)	(0.015)	(0.028)		

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel of the outcome. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Mean of the outcome is reported for children with a dyslexia score above cutoff. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

oral exam in math is one of the subjects that are randomly drawn from a pool of possible exam subjects. As a result, the sample size around the cutoff is substantively smaller than for the mandatory exams. From the table, we observe statistically significant results for the written exam in Danish as well as the subexams in spelling and reading. These results further confirm that there is a direct effect on the skills that dyslexic otherwise struggle with. However, we do not observe any effect on the exam in oral Danish. Additionally, we observe positive significant effects for all three subexams in math.

For the long run, we observe a positive effect of being dyslexic on continued enrollment

in the educational system two years after completion of compulsory schooling. Specifically, Appendix Table E2 demonstrates that they have a higher likelihood of transitioning from a no education trajectory to pursuing vocational education. This result indicates that children with dyslexia are able to overcome challenges and barriers associated with their condition if they receive adequate support.

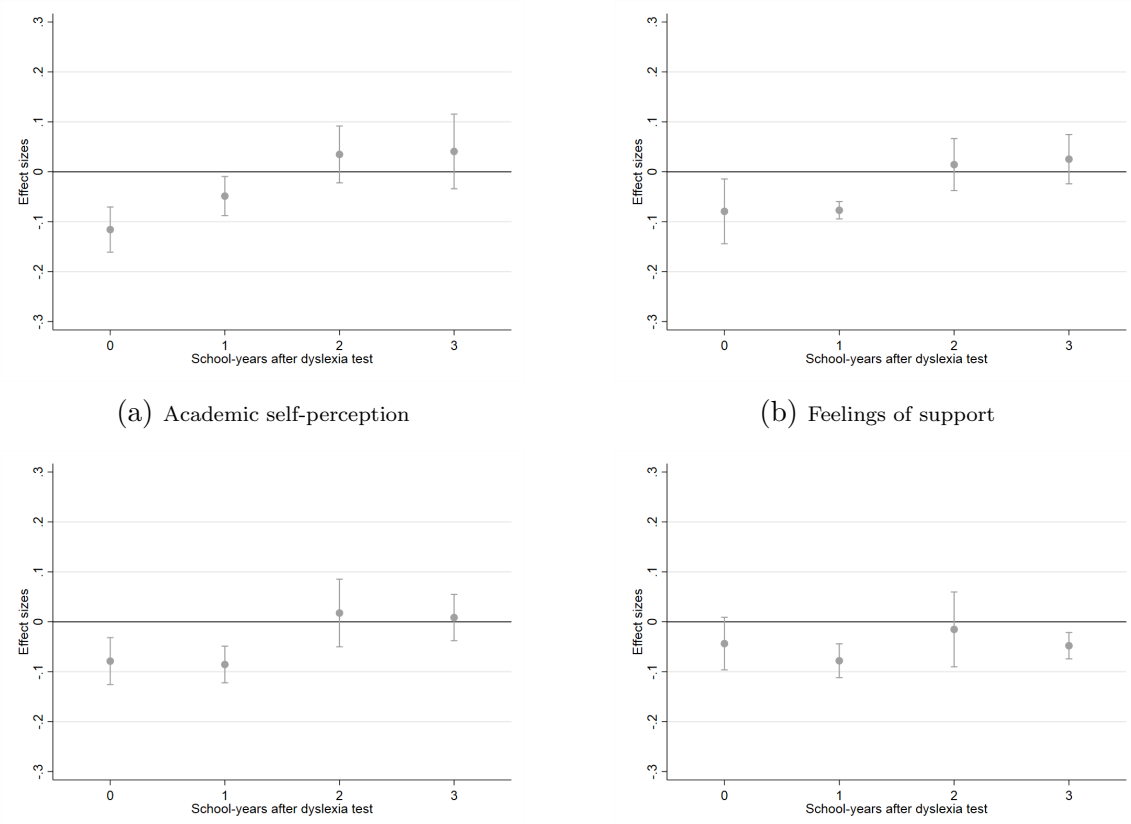
In column (3), we divide the reduced form coefficient in column (2) by the first stage coefficient in column (1) to obtain the LATE estimate. In all cases where the reduced-form estimates are statistically different from zero, the same holds true for the LATE estimates, which are mechanically larger.

In summary, our findings reveal short-term benefits following dyslexia diagnosis, with notable enhancements for reading and math test scores. Medium-term benefits persist particularly in Danish, driven by improvements in areas like spelling and reading where dyslexic students typically face challenges. These improvements extend to math and science, though English does not exhibit similar gains. Long-term observations reveal an increased propensity among dyslexic students to engage in vocational education after compulsory schooling, suggesting that with sufficient support, these students can successfully navigate challenges related to their condition.

C. Well-being outcomes

Next, we analyse the impact of being dyslexic and the targeted support on well-being. Specifically, we consider academic self-perception, feelings of support, academic well-being, and social well-being. These measures encapsulate a broad spectrum of experiences. Figure 4 shows each of the four well-being measures in the year of diagnosis and the following three years. Appendix Table D2 shows that there is no discontinuities across the cutoffs pre-diagnosis.

Interestingly, our findings suggest an initial drop in all four measures following the dyslexia diagnosis. This likely reflects the immediate impact of the diagnosis, which may cause some initial distress or confusion. However, the same measures show a steady upward trajectory following the initial drop until they after two years are insignificantly different from non-dyslexics. For academic self-perception in panel (a), this pattern could indicate a growth in the students' confidence in their academic capabilities over time as they become more comfortable with the digital tools and understand that they with the provided support indeed can thrive academically. While the effect may not reach statistical significance, there is a suggestive indication of a positive tendency. Similarly, for feelings of support in panel (b), children may gradually feel better supported and more engaged in their learning journey. Looking at academic well-being in panel (c) and social well-being in panel (d), it appears that a certain period of time is required before these measures return to their



(a) Academic self-perception

(b) Feelings of support

(c) Academic well-being

(d) Social well-being

Figure 4. : Well-being

The figure presents bias-corrected reduced-form estimates along with cluster-robust confidence intervals. The estimates are based on a sample of children with dyslexia scores within a 6-point average range around the cutoff. The figure focuses on well-being outcomes derived from the annual well-being survey conducted from grade 4 to 9.

normal levels. This delay may stem from the time required for students to process their diagnosis, adapt to new learning strategies, and fully utilize the support mechanisms in place.

Finally, in Appendix Figure E1, we investigate any changes in the frequency of absences, which could indicate changes in students' engagement and comfort with schooling post-diagnosis. We find no evidence of changes in absences during the first two years following diagnosis. However, there is some indication of an increase in absences in the third year, though we exercise caution in interpreting this estimate due to the relatively small number of observations.

D. Sensitivity checks

To evaluate the robustness of our findings, we conduct a comprehensive analysis to determine the sensitivity of our results to alternative specifications. This includes examining the impact of excluding covariates, varying bandwidth choices, and exploring different functional forms. By systematically testing these alternative scenarios, we aim to ensure that our main findings remain reliable and unaffected by specific modeling choices.

In our main specification, we carefully account for predetermined covariates by including them as control variables. To assess the robustness of our main findings regarding educational outcomes, we refer to Appendix Table F1, where we examine the impact of excluding these control variables in column (2). If the key assumption in our RD design is satisfied, omitting the control variables should not affect the point estimates but only impact precision. Notably, the point estimates in the first column, representing our main specification, remain largely unchanged even when the control variables are omitted in the second column. However, in the absence of control variables, the math score in the subsequent standardized test becomes statistically insignificant, although it still exhibits some indications of a positive effect. This sensitivity of the math score to the inclusion of control variables is not surprising given the significant imbalance observed in the math pre-test.

Next, we examine the robustness to alternative bandwidths and functional form. First, we in Appendix Table F1, column (3) allow the bandwidth to differ across outcomes using the CCT optimal bandwidth suggested by Calonico, Cattaneo and Titiunik (2014). Using this bandwidth seems to increase our estimates slightly throughout the considered outcomes. In column (4), we utilize a second-degree polynomial in dyslexia test score and find that the results remain consistent, with only slight changes observed due to the choice of functional form. This suggests that the findings are robust and not heavily influenced by the specific functional form employed. Second, in Appendix Figure F1, we present a comprehensive analysis using a wide range of bandwidths for both first- and second-degree polynomial specifications. The results from this figure demonstrate the robustness of our baseline estimates across various bandwidths and functional forms for the majority of educational outcomes. However, we observe that the initially significant effect on science exam grades diminishes when increasing bandwidth and is significant only in the close proximity to the cutoff.

In column (5), we perform an additional manipulation test to address a potential concern of heaping. Heaping refers to a situation where individuals manipulate their responses around the cutoff point, which could introduce bias if the heaping does not occur symmetrically around the cutoff. To examine the impact of heaping, we estimate 'doughnut' regressions by excluding children with dyslexia test scores at the cutoff. The results from

these 'doughnut' regressions are consistent with our main specification. This suggests that the main results are not primarily driven by heaping, providing further support for the robustness of our findings.

Our main specification employs a triangular kernel. In column (6), we assess the robustness of our findings by utilizing a rectangular kernel instead. Unlike the triangular kernel, the rectangular kernel assigns equal weights to all observations, without giving higher weight to those closer to the cutoff. The estimates remain highly robust when using the rectangular kernel. This finding further strengthens the validity of our results, demonstrating that they are not overly influenced by the specific choice of kernel function.

Finally, we perform a falsification test to assess the validity of our findings. We investigate whether similar discontinuities in educational outcomes exist at other points in the distribution of the dyslexia test. If the observed effects are truly driven by being diagnosed with dyslexia and receiving the dyslexia support package, we would not expect to observe consistent discontinuities in the outcomes at other cutoff values. To examine this, we analyze a range of cutoff values from -12 to 12, while keeping the bandwidth fixed at 6 points. The results, presented in Table F2, reveal that the dyslexia cutoff is unique, with the most substantial and statistically significant discontinuities found at the rescaled test score of 0. These findings further support the hypothesis that the observed effects are specifically linked to the dyslexia diagnosis and the targeted support.

In summary, the sensitivity analyses conducted provide strong evidence that the effects on educational performance can be attributed to being diagnosed with dyslexia and receiving dyslexia support. The robustness of the findings across various specifications, control variables, bandwidth choices, functional forms, and falsification tests underscores the causal link between the dyslexia diagnosis and the targeted support and the impact on educational outcomes.

V. Suggestive Mechanisms

Our main RD results provide evidence that being diagnosed with dyslexia leads to a substantial increase in academic performance without impacting well-being negatively in the long-run. We now turn to investigating whether certain characteristics or trajectories vary discontinuously at the cutoffs and whether this help to explain the effects on educational outcomes. In essence, we aim to identify groups of children who particularly benefit from receiving a dyslexia diagnosis and subsequently receiving the necessary support.

First, it is interesting to investigate whether there is a cumulative effect over the years. Appendix Table G1 shows the effects for the timing of the diagnosis and thus the duration for how long you have received dyslexia support. In terms of the reading score on the next standardized test, we find suggestive evidence of a significantly positive effect for grade

levels below 6 compared to above. However, while getting the diagnosis in Grade 3 or 4 seems to have a positive effect on the math final exam, the opposite seems to be the case for the Danish exam. However, the effect on GPA appears consistent across grade levels. In the long run, the largest effect is observed for early diagnoses.

Next, we explore heterogeneity by gender and socio-economic status (SES) in Appendix Table G2. In the short run, there seems to be a slightly higher effect on girls, which is persistent in the medium run for the exam in Danish. A similar effect seems to be present in terms of math. However, the GPA remains unchanged. In the long run, there appears to be a more substantial effect on boys in terms of remaining in an educational track. Regarding SES, we find a significant effect for low SES in the short run, particularly on the reading test. However, this effect does not persist into the medium run for the Danish exam. It is worth noting that the large effect for low SES may be influenced by a small sample size. Similarly, no significant differences in GPA are observed based on SES, but there appears to be a positive effect on math for low SES children. Interestingly, the long-term effects do not seem to be driven by low SES, but rather by high SES children.

We further examine whether the positive effects found in the standardized tests also carry over to the assessment by their teachers at the end of Grade 9. In Appendix Table G3, we report the grades that are given before the final exams by the children's subject-specific teacher. We find similar results as for the final exams for Danish, math, and science. However, there also seems to be a significant positive effect for English.

Finally, in the Appendix, Table G4, we thoroughly examine whether the outcomes might be influenced by a rise in enrollment in special education or private schools. Our findings indicate that there is no discernible impact in this regard.

VI. Conclusion

The inability of children to achieve proficient reading levels is a significant concern in education. This may be attributed to challenges in identifying children with reading disabilities, particularly dyslexia, and providing effective support systems. Dyslexia is a prevalent genetic neurological disorder that impairs reading, spelling, and writing abilities, affecting approximately 7% of children. These difficulties can lead to academic setbacks and impact overall well-being. However, evidence on dyslexia interventions remains limited, partly due to non-random diagnosis and selection biases.

To address these challenges, we examine the effects of diagnosing and supporting dyslexic children in compulsory schooling in Denmark. The implementation of a standardized testing scheme in 2015 provides an opportunity to use a regression discontinuity (RD) design to estimate the impact. By comparing children just above and below the dyslexia cutoff, this paper leverages a RD design to estimate effects for children on the margin of

diagnosis.

Our findings highlight the positive effect of dyslexia diagnosis on academic performance in the short term, specifically for reading test scores. These benefits continue into the medium term, particularly in areas where dyslexic students typically struggle, such as spelling and reading. Although English does not show similar gains, we find spillover effects to math and potentially science. Notably, in the long term, dyslexic students demonstrate an increased likelihood of pursuing vocational education after completing compulsory schooling, indicating their ability to overcome challenges associated with dyslexia with appropriate support. These results are robust to several sensitivity checks.

After the diagnosis, we observe a temporary decline in well-being across four distinct measures. However, encouragingly, our findings demonstrate that two years following the diagnosis, these measures return to their initial levels, indicating a recovery in well-being. Particularly for academic self-perception, there is even a positive tendency towards a potentially significant effect.

In summary, our findings demonstrate a significant increase on educational outcomes without negative effects on well-being in the long run for children diagnosed with dyslexia. These results have important implications for testing, diagnosing, and supporting children with reading disabilities. One important implication is the potential benefit of providing support to all children with reading difficulties, not just those diagnosed with dyslexia (Gibbs and Elliott, 2020). The findings suggest that even individuals who are marginally above the dyslexia diagnosis cutoff could benefit from access to digital tools or other support options for at least a shorter period of time. Additionally, this paper highlights the importance of standardized testing for dyslexia and subsequent provision of support for diagnosed children. Standardized testing can ensure accurate identification and enable effective targeted support to dyslexic children.

Overall, these findings emphasize the significance of early identification and adequate support for dyslexic children. By implementing appropriate support strategies, these students can achieve academic success and progress in the educational system.

REFERENCES

Andersen, Simon Calmar, Louise Beuchert, Helena Skyt Nielsen, and Mette Kjærgaard Thomsen. 2020*a*. "The Effect of Teacher's Aides in the Classroom: Evidence from a Randomized Trial." *Journal of the European Economic Association*, 18(1): 469–505.

Andersen, Simon Calmar, Miriam Gensowski, Steven G. Ludeke, and Oliver P. John. 2020*b*. "A stable relationship between personality and academic performance from childhood through adolescence. An original study and replication in hundred-thousand-person samples." *Journal of Personality*, 88(5): 925–939.

Beuchert, Louise V, and Anne B Nandrup. 2018. "The Danish National Tests at a Glance." *Nationaløkonomisk Tidsskrift*, 2.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2019. "Regression Discontinuity Designs Using Covariates." *The Review of Economics and Statistics*, 101(3): 442–451.

Currie, Janet, and Mark Stabile. 2006. "Child mental health and human capital accumulation: The case of ADHD." *Journal of Health Economics*, 25(6): 1094–1118.

Doust, C., P. Fontanillas, E. Eising, S. D. Gordon, Z. Wang, G. Alagöz, B. Molz, 23andMe Research Team, Quantitative Trait Working Group of the GenLang Consortium, B. Pourcain, C. Francks, R. E. Marioni, J. Zhaa, S. Paracchini, J. B. Talcott, A. P. Monaco, J. F. Stein, J. R. Gruen, R. K. Olson, E. G. Willcutt, J. C. DeFries, B. F. Pennington, S. D. Smith, M. J. Wright, N. G. Martin, A. Auton, T. C. Bates, S. E. Fisher, and M. Luciano. 2022. "Discovery of 42 genome-wide significant loci associated with dyslexia." *Nature Genetics*, 54: 1621–1629.

Démonet, Jean-François, Margot J Taylor, and Yves Chaix. 2004. "Developmental dyslexia." *The Lancet*, 363: 1451–1460.

Erbeli, Florina, Marianne Rice, and Silvia Paracchini. 2022. "Insights into Dyslexia Genetics Research from the Last Two Decades." *Brain Sciences*, 12(1): 27.

Ferman, Bruno, and Luiz Felipe Fontes. 2022. "Assessing knowledge or classroom behavior? Evidence of teachers' grading bias." *Journal of Public Economics*, 216: 104773.

Fletcher, Jason M. 2014. "THE EFFECTS OF CHILDHOOD ADHD ON ADULT LABOR MARKET OUTCOMES." *Health Economics*, 23(2): 159–181.

Gabrieli, John D. E. 2009. "Dyslexia: A New Synergy Between Education and Cognitive Neuroscience." *Science*, 325(5938): 280–283.

Gibbs, Simon J., and Julian G. Elliott. 2020. "The dyslexia debate: life without the label." *Oxford Review of Education*, 46(4): 487–500.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2018. "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking." *Journal of Political Economy*, 126(S1): S197–S246.

Hoeft, Fumiko, Bruce D. McCandliss, Jessica M. Black, Alexander Gantman, Nahal Zakerani, Charles Hulme, Heikki Lyytinen, Susan Whitfield-Gabrieli, Gary H. Glover, Allan L. Reiss, and John D. E. Gabrieli. 2011. "Neural systems predicting long-term outcome in dyslexia." *Proceedings of the National Academy of Sciences*, 108(1): 361–366.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Kale, Sirin. 2020. "The battle over dyslexia." *The Guardian*.

Lee, David S., and David Card. 2008. "Regression discontinuity inference with specification error." *Journal of Econometrics*, 142(2): 655–674. The regression discontinuity design: Theory and applications.

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

Lyon, G. Reid, Sally E. Shaywitz, and Bennett A. Shaywitz. 2003. "A definition of dyslexia." *Annals of Dyslexia*, 53: 1–14.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714. The regression discontinuity design: Theory and applications.

Peterson, Robin L., and Bruce F. Pennington. 2012. "Developmental dyslexia." *The Lancet*, 379(9830): 1997–2007.

Shaywitz, Bennett A., and Sally E. Shaywitz. 2020. "The American experience: towards a 21st century definition of dyslexia." *Oxford Review of Education*, 46(4): 454–471.

Shaywitz, Sally E. 1998. "Dyslexia." *New England Journal of Medicine*, 338(5): 307–312.

Snowling, Margaret J., Charles Hulme, and Kate Nation. 2020. "Defining and understanding dyslexia: past, present and future." *Oxford Review of Education*, 46(4): 501–513.

Toffalini, Enrico, David Giofre, Massimiliano Pastore, Barbara Carretti, Federica Fraccadori, and Denes Szucs. 2021. "Dyslexia treatment studies: A systematic review and suggestions on testing treatment efficacy with small effects and small samples." *Behavior Research Methods*, 53: 1954–1972.

Wagner, Richard K., Fotena A. Zirps, Ashley A. Edwards, Sarah G. Wood, Rachel E. Joyner, Betsy J. Becker, Guangyun Liu, and Bethany Beal. 2020. "The Prevalence of Dyslexia: A New Approach to Its Estimation." *Journal of Learning Disabilities*, 53(5): 354–365. PMID: 32452713.

Tests and Surveys

Every year starting from 2010, reading tests are administered to students in Grades 2, 4, 6, and 8, while math tests are given to students in Grades 3, 6, and, since 2018, Grade 8. In addition, English tests are conducted for students in Grade 7, and science tests are administered to students in grade 8 during the spring.

Table A1—: Frequency of Tests

Grade	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Reading test			X		X		X		X	
Math test				X			X			
Well-being survey					X	X	X	X	X	X

Note: The table shows the frequency of tests and well-being surveys in compulsory schooling in Denmark.

Test Composition

The Danish dyslexia test consists of three distinct subtests. The first subtest involves the student listening to a pseudoword and selecting the correct spelling from a set of five options. To accomplish this, the student must decode the potential pseudoword spellings and match them with the one they heard. A time limit of 5 minutes is given for completing as many of the 40 items as possible.

The second subtest is a phonological choice test, requiring the student to determine which pseudoword, out of five options, sounds like a real word when pronounced. The student is allotted 7 minutes to complete as many of the 44 items as they can.

The third subtest assesses vocabulary and includes 15 items, but it does not provide a subscore for the overall results of the test. Its purpose is to identify students whose Danish vocabulary might be inadequate for the two first subtests to be reliable. If the child gets less than 10 items correct, the reliability needs to be questioned.

A standard subscore is automatically calculated for each of the two first subtests to evaluate efficiency. This subscore considers both accuracy (adjusted for guessing) and time (adjusted for extremes). Specifically, they are calculated in the following way:

$$\frac{\text{number of correct answers} - \frac{\text{number of wrong answers}}{\text{number of wrong options}}}{\text{number of answers} * \text{median speed of answers}}$$

The overall standard score is then obtained by taking the average of these two subscores.

Sample Selection

Table C1—: Sample Selection

National dyslexia test between 2014/2015 and 2018/2019	**67,533**
- Enrolled at private and public schooling	65,151
- In regular classroom	60,952
- Tested during the spring (validated test time)	43,086
- No membership of dyslexia database prior to dyslexia test	35,665
- Not recently diagnosed with mental disorder	34,425
- Having at least one primary outcome	31,610

RD DESIGN

Table D1—: Bandwidth Selection

	Bandwidth estimate (1)	Bandwidth bias (2)
Short run academic outcomes		
- First reading score (std.)	4.579	8.505
- Math score (std.)	7.620	14.250
Exam outcomes		
- GPA (std.)	4.770	9.619
- Danish (std.)	4.996	9.255
- Math (std.)	4.629	9.415
- English (std.)	4.340	8.458
- Science (std.)	6.204	11.448
No education post schooling		
- Two years after (1/0)	5.594	11.499

Note: Each cell indicates the optimal bandwidth using the methodology suggested by Calonico et. al (2014,2019). Column 1 reports the optimal bandwidth for point estimators using a first-degree polynomial and column 2 show optimal bandwidth for inference using a second-degree polynomial.

Table D2—: Distribution of Covariates Across the Dyslexia Cutoff

	Estimate (1)	BC estimate (2)	Robust S.E. (3)	Mean (4)	N (5)
A. Survey items at dyslexic test					
- Difficult to learn reading and spelling? (1/0)	0.010	[0.008]	(0.008)	0.773	16,630
- Receive additional help in reading? (1/0)	0.004	[-0.002]	(0.023)	0.365	16,630
- Previously received additional help? (1/0)	0.000	[-0.005]	(0.021)	0.610	16,630
- Do you have difficulty reading? (1/0)	*0.019*	*[0.018]*	(0.010)	0.614	16,630
- Do you have difficulty spelling? (1/0)	-0.006	[-0.011]	(0.007)	0.792	16,630
- Do you think that you are dyslexic? (1/0)	-0.002	[-0.005]	(0.012)	0.536	16,630
- Teacher told you that you are dyslexic? (1/0)	0.005	[0.010]	(0.008)	0.062	16,630
- Do you have tools that you peers don't? (1/0)	-0.019	[-0.028]	(0.018)	0.303	16,630
B. Child characteristics					
- Boy (1/0)	-0.025	[-0.022]	(0.015)	0.568	16,630
- Western etnicity (1/0)	0.001	[0.001]	(0.004)	0.011	16,630
- Non-western etnicity (1/0)	0.007	[0.008]	(0.007)	0.053	16,630
- Living with both parents (1/0)	**-0.030**	**[-0.037]**	(0.007)	0.693	16,630
- School starting age (years)	-0.002	[0.000]	(0.005)	6.143	16,000
- Privat schooling (1/0)	-0.006	[-0.008]	(0.007)	0.101	16,630
- Special needs teaching (1/0)	-0.000	[-0.000]	(0.001)	0.002	16,519
- Special needs classroom (1/0)	0.000	[-0.000]	(0.001)	0.002	16,519
- Psychiatric diagnosis (1/0)	-0.001	[-0.002]	(0.002)	0.014	16,630
- Reading score (std.)	-0.020	[-0.035]	(0.027)	-0.688	14,503
- Math score (std.)	**-0.034**	**[-0.032]**	(0.016)	-0.423	13,373
- Conscientiousness (std.)	0.008	[-0.005]	(0.056)	-0.210	5,944
- Agreeableness (std.)	*0.093*	*[0.110]*	(0.066)	-0.114	6,054
- Emotional stability (std.)	**-0.071**	**[-0.079]**	(0.031)	-0.034	5,806
- Academic self-preception (std.)	-0.032	[-0.032]	(0.037)	-0.276	5,319
- Academic well-being (std.)	-0.007	[-0.015]	(0.051)	-0.273	6,143
- Social well-being (std.)	-0.029	[-0.019]	(0.023)	-0.017	6,164
- Support and inspriation (std.)	0.012	[0.020]	(0.024)	0.043	6,149
- Absence (percent)	-0.004	[0.009]	(0.050)	1.626	14,844
- Age at dyslexia test date	-0.071	[-0.066]	(0.057)	11.970	16,630
B. Mother's characteristics					
- Age at birth (years)	0.151	[0.174]	(0.148)	30.361	16,594
- Employed (1/0)	-0.005	[-0.003]	(0.011)	0.748	16,591
- High school or less (1/0)	-0.005	[-0.002]	(0.010)	0.202	16,630
- Education length (years)	0.072	[0.071]	(0.068)	14.346	16,417
- Work experience (years)	0.001	[0.048]	(0.242)	14.412	16,511
- Income (1,000DKK)	*3.878*	*[7.327]*	(3.884)	291.570	16,537
C. Father's characteristics					
- Age at birth (years)	0.021	[-0.021]	(0.144)	32.898	16,424
- Employed (1/0)	**-0.018**	**[-0.021]**	(0.007)	0.850	16,443
- High school or less (1/0)	0.007	[0.017]	(0.011)	0.250	16,630
- Education length (years)	-0.032	[-0.064]	(0.065)	14.110	16,000
- Work experience (years)	-0.455	[-0.635]	(0.439)	18.841	16,153
- Income (1,000DKK)	**-22.612**	**[-29.045]**	(12.224)	416.584	16,263

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Columns 1-5 report reduced-form results from separate local-linear regressions similar to the main equation with outcome indicated in the row. All regressions use a triangular kernel. Column 1 reports conventional estimates, Column 2 bias-corrected estimates, Column 3 robust standard errors, Column 4 the mean of the dependent variables for observations to the right of the cutoff, and Column 5 number of observations. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

Results

Table E1—: Subexams for the Final Exams

	First stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean (4)	N (5)
Danish					
Written	**0.527**	**0.079**	**0.149**		
	[0.530]	**[0.101]**	**[0.191]**	-0.545	10,768
	(0.008)	(0.024)	(0.046)		
Spelling	**0.527**	**0.122**	**0.231**		
	[0.530]	**[0.129]**	**[0.244]**	-0.754	10,801
	(0.008)	(0.016)	(0.030)		
Oral	**0.526**	0.005	0.010		
	[0.529]	[0.001]	[0.003]	-0.421	10,747
	(0.008)	(0.013)	(0.024)		
Reading	**0.527**	**0.124**	**0.233**		
	[0.530]	**[0.129]**	**[0.244]**	-0.586	10,803
	(0.008)	(0.011)	(0.022)		
Math					
Basic skills	**0.527**	**0.069**	**0.131**		
	[0.530]	**[0.078]**	**[0.147]**	-0.434	10,800
	(0.008)	(0.016)	(0.030)		
Problem-solving	**0.527**	**0.067**	**0.126**		
	[0.530]	**[0.075]**	**[0.143]**	-0.390	10,769
	(0.008)	(0.017)	(0.032)		
Oral	**0.513**	**0.075**	**0.144**		
	[0.518]	**[0.102]**	**[0.198]**	-0.361	2,106
	(0.018)	(0.037)	(0.071)		

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel of the covariate. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Mean of the covariate is reported for children with a dyslexia score above cutoff. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

Table E2—: Education Trajectories after Completion of Compulsory Schooling

	Estimate (1)	BC estimate (2)	Robust S.E. (3)	Mean (4)	N (5)
Two years after primary schooling					
- No education (1/0)	**-0.040**	**[-0.047]**	(0.015)	0.140	8,143
- High-school (1/0)	0.004	[0.010]	(0.017)	0.526	8,143
- Vocational training (1/0)	*0.036*	*[0.038]*	(0.020)	0.334	8,143

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel of the covariate. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Mean of the covariate is reported for children with a dyslexia score above cutoff. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

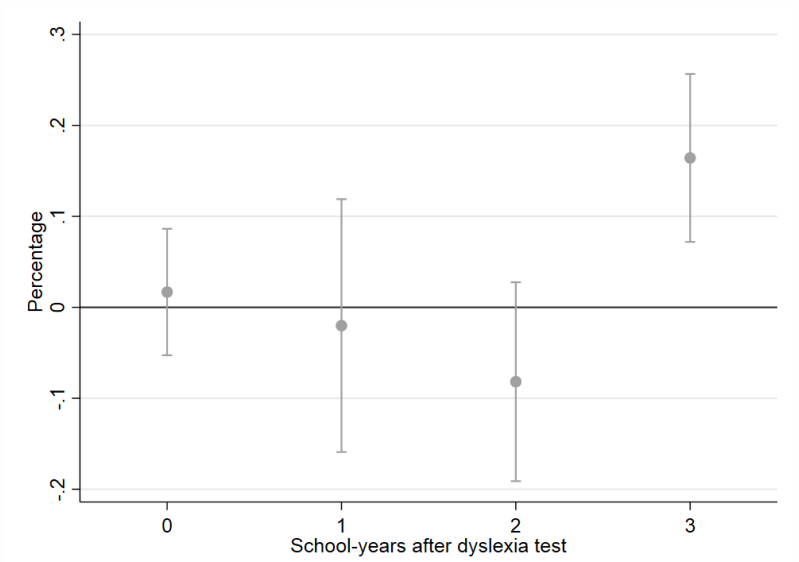


Figure E1. : Absence

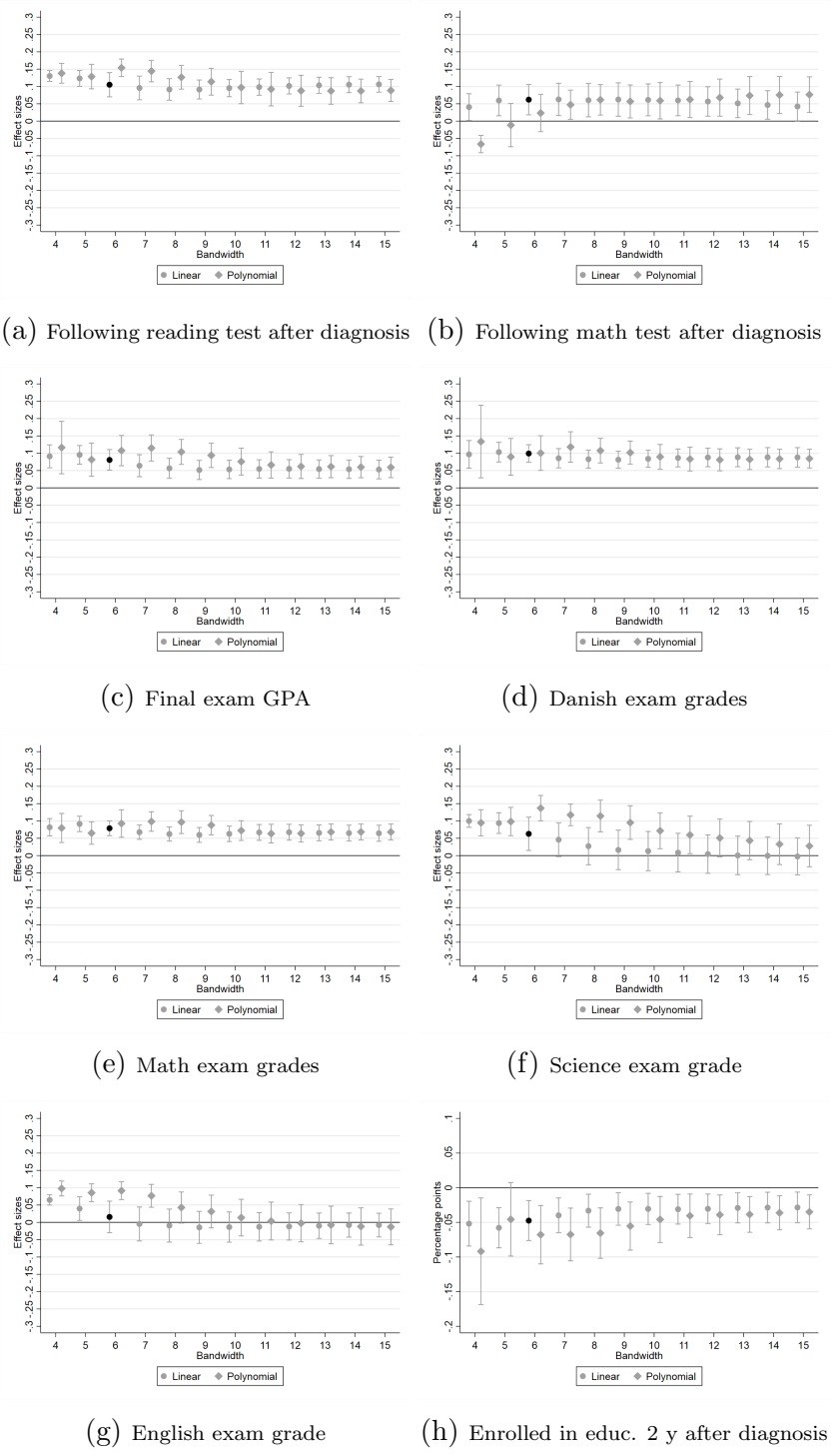
The figure presents bias-corrected reduced-form estimates along with cluster-robust confidence intervals. The estimates are based on a sample of children with dyslexia scores within a 6-point average range around the cutoff.

Sensitivity Checks

Table F1—: Sensitivity Checks

	Main spec. (1)	No controls (2)	CCT opt. BW (3)	2-degree poly. (4)	Donut reg. (5)	Rec. kernel (6)
Reading score	**0.101**	**0.091**	**0.111**	**0.133**	**0.095**	**0.110**
	[0.105]	**[0.091]**	**[0.117]**	**[0.149]**	**[0.099]**	**[0.114]**
	(0.018)	-0.018	(0.016)	(0.011)	(0.020)	(0.021)
Effective # of Obs.	13,283	13,283	11,196	13,283	11,983	11,196
Math score	**0.055**	0.036	**0.055**	0.012	0.026	**0.054**
	[0.062]	[0.041]	**[0.062]**	[-0.003]	[0.033]	**[0.067]**
	(0.022)	-0.037	(0.022)	(0.029)	(0.024)	(0.032)
Effective # of Obs.	12,068	12,068	13,712	12,068	10,880	8,073
GPA	**0.068**	**0.080**	**0.072**	**0.091**	**0.067**	**0.074**
	[0.081]	**[0.094]**	**[0.081]**	**[0.102]**	**[0.084]**	**[0.084]**
	(0.015)	-0.027	(0.016)	(0.019)	(0.023)	(0.015)
Effective # of Obs.	10,879	10,879	10,879	10,879	9,806	7,252
Danish	**0.092**	**0.115**	**0.093**	**0.108**	**0.099**	**0.097**
	[0.099]	**[0.123]**	**[0.100]**	**[0.111]**	**[0.117]**	**[0.107]**
	(0.013)	-0.027	(0.013)	(0.019)	(0.019)	(0.014)
Effective # of Obs.	10,825	10,825	10,825	13,736	9,756	9,131
Math	**0.071**	**0.063**	**0.072**	**0.093**	**0.068**	**0.074**
	[0.079]	**[0.073]**	**[0.077]**	**[0.094]**	**[0.082]**	**[0.081]**
	(0.011)	-0.02	(0.011)	(0.015)	(0.023)	(0.012)
Effective # of Obs.	10,810	10,810	10,810	13,704	9,747	7,212
English	0.005	-0.001	**0.025**	**0.070**	-0.016	0.011
	[0.016]	[0.011]	**[0.040]**	**[0.089]**	[-0.009]	[0.022]
	(0.023)	-0.019	(0.018)	(0.011)	(0.027)	(0.024)
Effective # of Obs.	10,617	10,617	8,966	10,617	9,567	8,966
Science	**0.040**	**0.052**	**0.066**	**0.099**	0.007	0.007
	[0.063]	**[0.077]**	**[0.093]**	**[0.120]**	[0.021]	[0.016]
	(0.024)	-0.031	(0.014)	(0.019)	(0.023)	(0.029)
Effective # of Obs.	10,465	10,465	8,840	10,465	9,432	11,942
No educ. 2 years after	**-0.040**	**-0.041**	**-0.043**	**-0.060**	**-0.042**	**-0.048**
	[-0.047]	**[-0.048]**	**[-0.049]**	**[-0.066]**	**[-0.048]**	**[-0.050]**
	(0.015)	-0.016	(0.015)	(0.019)	(0.018)	(0.024)
Effective # of Obs.	8,143	8,143	8,143	10,423	7,325	6,844

Note: Each cell reports the reduced-form coefficient from a separate local-linear regression with a triangular kernel of the outcome listed in the row. Additional covariates included in column 2 are show in the table D2. Column 3 and 4 used CCT optimal BW. Column 5 excluded observations at the margin. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).



(a) Following reading test after diagnosis (b) Following math test after diagnosis

(c) Final exam GPA (d) Danish exam grades

(e) Math exam grades (f) Science exam grade

(g) English exam grade (h) Enrolled in educ. 2 y after diagnosis

Figure F1. : Bandwidth and functional form sensitivity checks

The figure demonstrates the robustness of the main results to alternative bandwidth and functional form specifications. The dots indicate the bias-corrected estimate obtained using a first-degree polynomial, while the diamonds represent estimates derived from a second-degree polynomial. The vertical segments depict the corresponding 95% robust confidence intervals. The black dot represents the baseline results.

Table F2—: Discontinuities in outcomes at other points in the dyslexia test distribution

	-12 (1)	-9 (2)	-6 (3)	-3 (4)	0 (5)	3 (6)	6 (7)	9 (8)	12 (9)
Reading score	-0.025	*0.057*	-0.024	-0.032	**0.101**	-0.037	**-0.030**	**0.093**	0.008
	(0.029)	(0.030)	(0.027)	(0.028)	(0.018)	(0.024)	(0.015)	(0.029)	(0.053)
Effective # of Obs.	5,226	8,317	11,329	13,149	13,283	11,477	8,641	5,755	3,371
Math score	-0.038	0.023	-0.001	*-0.041*	**0.056**	-0.026	0.001	0.023	0.035
	(0.056)	(0.041)	(0.045)	(0.022)	(0.022)	(0.023)	(0.016)	(0.039)	(0.050)
Effective # of Obs.	4,556	7,317	10,110	11,835	12,068	10,459	7,881	5,255	3,066
GPA	-0.041	-0.004	0.012	**-0.049**	**0.069**	*-0.040*	0.016	*-0.039*	-0.019
	(0.027)	(0.026)	(0.016)	(0.014)	(0.015)	(0.021)	(0.018)	(0.021)	(0.019)
Effective # of Obs.	4,646	6,975	9,234	10,728	10,879	9,523	7,438	5,126	3,250
Danish	-0.031	-0.021	-0.012	**-0.044**	**0.093**	*-0.051*	**0.043**	-0.002	0.031
	(0.034)	(0.023)	(0.025)	(0.014)	(0.013)	(0.030)	(0.020)	(0.023)	(0.029)
Effective # of Obs.	4,613	6,924	9,181	10,674	10,825	9,480	7,398	5,101	3,235
Math	**-0.065**	-0.003	*0.039*	**-0.043**	**0.071**	-0.040	-0.005	-0.003	-0.067
	(0.027)	(0.038)	(0.024)	(0.017)	(0.011)	(0.026)	(0.028)	(0.023)	(0.045)
Effective # of Obs.	4,595	6,906	9,157	10,647	10,810	9,457	7,385	5,089	3,226
English	*-0.066*	0.001	-0.040	0.006	0.005	0.008	0.011	-0.028	-0.011
	(0.037)	(0.027)	(0.029)	(0.026)	(0.023)	(0.014)	(0.017)	(0.038)	(0.049)
Effective # of Obs.	4,424	6,682	8,907	10,414	10,617	9,322	7,283	5,026	3,186
Science	*-0.047*	**0.070**	*0.070*	-0.035	0.040	-0.010	**-0.032**	**-0.051**	-0.018
	(0.027)	(0.027)	(0.041)	(0.025)	(0.025)	(0.023)	(0.011)	(0.026)	(0.026)
Effective # of Obs.	4,330	6,557	8,744	10,251	10,465	9,191	7,196	4,962	3,143
No educ. 2 years after	**0.041**	*-0.027*	-0.003	**0.044**	**-0.040**	-0.003	-0.008	0.016	0.009
	(0.013)	(0.015)	(0.017)	(0.019)	(0.015)	(0.014)	(0.015)	(0.017)	(0.029)
Effective # of Obs.	3,702	5,432	7,035	8,085	8,143	7,168	5,608	3,902	2,509

Note: Sample of children with dyslexia score within a 6 point average around the different cutoffs indicated in the column headings. We report the conventional reduced-form estimates and robust standard errors in parentes from separate local-linear regressions. All regressions use a triangular kernel. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals (for more information, see Calonico et al., 2014; 2019).

Table G1—: Heterogeneous treatment effect by grade

	Estimate (1)	BC estimate (2)	Robust S.E. (3)	Mean (4)	N (5)
A. Reading score					
- Grade 3 & 4	**0.089**	[**0.093**]	(0.021)	-0.699	7,531
- Grade 5 & 6	**0.124**	[**0.132**]	(0.062)	-0.615	4,032
- Grade 7 & 8	0.055	[0.043]	(0.042)	-0.641	1,720
B. GPA					
- Grade 3 & 4	**0.065**	[**0.072**]	(0.031)	-0.617	2,845
- Grade 5 & 6	**0.077**	[**0.101**]	(0.024)	-0.537	4,251
- Grade 7, 8, & 9	**0.071**	[**0.075**]	(0.037)	-0.601	3,783
C. Danish					
- Grade 3 & 4	0.036	[0.035]	(0.032)	-0.695	2,820
- Grade 5 & 6	**0.112**	[**0.129**]	(0.022)	-0.630	4,235
- Grade 7, 8, & 9	**0.120**	[**0.123**]	(0.033)	-0.682	3,770
D. Math					
- Grade 3 & 4	**0.095**	[**0.091**]	(0.044)	-0.471	2,821
- Grade 5 & 6	**0.077**	[**0.086**]	(0.024)	-0.398	4,224
- Grade 7, 8, & 9	**0.061**	[**0.075**]	(0.037)	-0.433	3,765
E. No education two years after compulsory schooling					
- Grade 3 & 4	**-0.093**	[**-0.117**]	(0.028)	0.158	1,463
- Grade 5 & 6	-0.032	[-0.034]	(0.023)	0.145	2,901
- Grade 7, 8, & 9	**-0.038**	[**-0.043**]	(0.016)	0.129	3,779

Note: Sample of children with dyslexia score within a 6 point average around the cutoff divide into grade groups. Columns 1-5 report reduced-form results from separate local-linear regressions similar to the main equation with outcome indicated in the row. All regressions use a triangular kernel. Column 1 reports conventional estimates, Column 2 bias-corrected estimates, Column 3 robust standard errors, Column 4 the mean of the dependent variables for observations to the right of the cutoff, and Column 5 number of observations. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

Table G2—: Heterogeneous treatment effect by Gender and SES

		Sex		SES	
	All	Boy	Girls	Low	High
	(1)	(2)	(3)	(4)	(5)
Reading score	**0.101**	**0.085**	**0.122**	**0.235**	**0.071**
	[0.106]	**[0.078]**	**[0.141]**	**[0.252]**	**[0.075]**
	(0.018)	(0.020)	(0.047)	(0.090)	(0.011)
Effective # of Obs.	13,283	7,413	5,870	2,789	10,494
GPA	**0.069**	**0.068**	**0.068**	**0.079**	**0.066**
	[0.082]	**[0.080]**	**[0.084]**	**[0.099]**	**[0.076]**
	(0.015)	(0.013)	(0.024)	(0.044)	(0.012)
Effective # of Obs.	10,879	6,063	4,816	2,202	8,677
Danish	**0.093**	**0.077**	**0.111**	**0.074**	**0.099**
	[0.100]	**[0.084]**	**[0.119]**	**[0.082]**	**[0.107]**
	(0.013)	(0.016)	(0.016)	(0.039)	(0.012)
Effective # of Obs.	10,825	6,036	4,789	2,181	8,644
Math	**0.071**	**0.049**	**0.096**	**0.095**	**0.064**
	[0.080]	**[0.056]**	**[0.108]**	**[0.121]**	**[0.067]**
	(0.011)	(0.021)	(0.024)	(0.051)	(0.012)
Effective # of Obs.	10,810	6,035	4,775	2,179	8,631
No educ. 2 y. after	**-0.040**	**-0.051**	**-0.029**	0.003	**-0.050**
	[-0.047]	**[-0.061]**	**[-0.030]**	[0.017]	**[-0.062]**
	(0.015)	(0.026)	(0.015)	(0.036)	(0.021)
Effective # of Obs.	8,143	4,571	3,572	1,682	6,461

Note: The analysis is based on a sample of children with dyslexia scores within a 6-point average range around the cutoff. The sample is further stratified by sex and socioeconomic status (SES), specifically based on the mother's educational level, distinguishing between high school or below and those with higher educational attainment. Columns 1-5 report reduced-form results from separate local-linear regressions similar to the main equation with outcome indicated in the row. All regressions use a triangular kernel. Column 1 reports conventional estimates, Column 2 bias-corrected estimates, Column 3 robust standard errors, Column 4 the mean of the dependent variables for observations to the right of the cutoff, and Column 5 number of observations. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

Table G3—: Year-grades

	First-stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean outcome (4)	Effective obs. (5)
Danish	**0.524**	**0.060**	**0.114**	-0.637	10,900
	[0.527]	**[0.066]**	**[0.127]**		
	(0.008)	(0.018)	(0.034)		
Math	**0.525**	**0.055**	**0.104**	-0.412	10,897
	[0.528]	**[0.061]**	**[0.116]**		
	(0.008)	(0.008)	(0.016)		
English	**0.526**	**0.016**	**0.030**	-0.608	10,810
	[0.529]	**[0.028]**	**[0.054]**		
	(0.008)	(0.014)	(0.027)		
Science	**0.529**	**0.031**	**0.059**	-0.341	10,580
	[0.532]	**[0.054]**	**[0.101]**		
	(0.008)	(0.018)	(0.033)		

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel of the outcome. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Mean of the outcome is reported for children with a dyslexia score above cutoff. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).

Table G4—: Enrollment into special education and private schooling

	Estimate (1)	BC estimate (2)	Robust S.E. (3)	Mean (4)	N (5)
A. Special education					
- year 1 (1/0)	-0.000	[0.001]	(0.002)	0.005	11,983
- year 2 (1/0)	-0.001	[0.002]	(0.005)	0.013	7,488
- year 3 (1/0)	-0.011	[-0.010]	(0.015)	0.031	4,051
B. Private schooling					
- year 1 (1/0)	0.005	[0.005]	(0.004)	0.118	12,104
- year 2 (1/0)	0.002	[0.007]	(0.012)	0.136	7,954
- year 3 (1/0)	**-0.024**	**[-0.027]**	(0.009)	0.144	4,820

Note: Sample of children with dyslexia score within a 6 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel of the outcome. Bias-corrected estimates are listed in square brackets and robust standard errors in brackets below. Mean of the outcome is reported for children with a dyslexia score above cutoff. Bold (italic) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (for more information, see Calonico et al., 2014; 2019).