Marginal Diagnoses, Lasting Effects: Evidence from a Nationwide Dyslexia Test

By Søren Albeck Nielsen & Kristoffer Ibsen*

This paper provides the first causal evidence on the effects of diagnosing dyslexia and providing mandated educational support during compulsory schooling. Dyslexia affects roughly 7 percent of all children, yet causal evidence on the consequences of formal diagnosis remains limited, particularly given selection problems arising from inconsistent and subjective diagnostic practices. Exploiting a nationwide standardized test implemented in Denmark in 2015, we use a regression discontinuity design to identify these effects for children at the margin of diagnosis. We find large and persistent academic gains, with diagnosed students achieving higher reading performance and school-leaving GPA, and these gains translate into meaningful long-run outcomes. Diagnosis and mandated support raise the likelihood of completing an academic high school track by 16 percent and lower the likelihood of being not in education, employment, or training (NEET) at age 20 by 49 percent. Well-being initially declines after diagnosis, but this reverses within two years, with significant gains in school well-being and academic self-perception four years after. Long-run effects are larger for students identified earlier in compulsory schooling, consistent with dynamic complementarities between the timing of diagnosis and the returns to support. Our findings suggest that timely diagnosis and mandated support can generate substantial and lasting improvements in children's human capital accumulation and well-being.
JEL: H52, I21
Keywords: Dyslexia, diagnosis, 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)

Reading proficiency plays a central role in human capital accumulation. Yet a substantial share of students in developed countries fail to reach basic reading benchmarks. In the

* TrygFonden's Centre for Child Research & Department of Economics and Business Economics, Aarhus University. An earlier version of the draft was previously circulated with the title: "Unlocking Potential: The Impact of Dyslexia Diagnosis and Support on Educational Outcomes and Well-being". We gratefully acknowledge helpful comments and discussions from Eric Bettinger, Elena Mattana, Michael Rosholm, Alexander Koch as well as conference and seminar participants at various universities and institutions.

PISA 2025 assessment, 31 percent of students in OECD countries performed below the minimum proficiency level in reading (OECD, 2026).[1] Left unaddressed, a child's reading difficulties tend to persist over time (Heckman and Mosso, 2014; Houmark et al., 2025), limiting which subjects a student can engage with, their confidence in their own abilities, and ultimately the educational trajectory they can pursue. For some of these children, the underlying reason is a reading disability. This paper provides the first causal evidence on how the diagnosis and mandated support of children with dyslexia, the most common reading disability, affect their educational outcomes and well-being.

Dyslexia is a neurodevelopmental disorder with substantial genetic components that impairs individuals' reading, spelling, and writing abilities. Rather than being a binary condition, dyslexia exists on a continuum of severity, and diagnosis typically involves applying a cutoff to a continuous measure of reading ability (Gabrieli, 2009). It is estimated to affect approximately 7 percent of all children, corresponding to one to two students in a typical classroom, though estimates range up to 17.5 percent depending on the diagnostic criteria applied (Shaywitz, 1998; Peterson and Pennington, 2012; Yang et al., 2022). Despite dyslexia's notable prevalence, there is a remarkable scarcity of evidence on the outcomes of receiving a dyslexia diagnosis compared with remaining undiagnosed. 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).

This gap partly reflects a deeper identification problem. 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. Even trained professionals facing similar cases vary systematically, not only in where they set the diagnostic threshold, but in their underlying diagnostic skill (Chan, Gentzkow and Yu, 2022), and the same child can receive different results across repeated tests using different criteria (Erbeli, Rice and Paracchini, 2022). Beyond this, dyslexia assessments can be costly and time-consuming, and identification depends in part on which school a child attends (Daniel et al., 2026), so that children with similar underlying dyslexia severity may have systematically different chances of being diagnosed. Consequently, a naive comparison of diagnosed and undiagnosed children is unlikely to identify the causal effect of diagnosis alone, since it would also reflect these sources of measurement inconsistency and selection.

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

[1]This share has risen steadily. Among the 35 OECD countries with comparable data since 2012, the share of students below the minimum proficiency level in reading increased from 19 percent in 2012 to 30 percent in 2025 (OECD, 2026).

schooling from age 6 to 15. The test is individual, computer-based, and conducted under the supervision of a trained instructor with a teaching background. This implementation eliminates the reliance on subjective evaluations and ensures a uniform procedure of diagnosis across the student population. Upon diagnosis, students become legally entitled to mandated support, which typically includes digital tools such as text-to-speech software, individualized teaching adaptations, and modified exam conditions such as extended time and continued access to these digital tools. Leveraging the standardized dyslexia test, we adopt a regression discontinuity (RD) design to identify the causal effects of diagnosis and mandated support. Importantly, children who score just above or just below the diagnostic cutoff on a given test day are similar in observed characteristics and, plausibly, in unobserved ones as well, so that assignment to diagnosis near the cutoff is as good as random, allowing us to estimate causal effects for students at the margin of diagnosis. Because some children obtain a diagnosis outside the standardized testing system, we complement our sharp RD estimates, which identify the intent-to-treat (ITT) effect of crossing the diagnostic threshold, with a fuzzy RD design that yields the local average treatment effect (LATE).

Our first empirical finding is that receiving a dyslexia diagnosis and subsequent support causally improves outcomes across multiple stages of the education trajectory. In the short run, diagnosed students achieve higher reading performance, increasing test scores by 0.097 SD. These gains extend to school-leaving exam GPA, which rises by 0.063 SD, an effect driven primarily by improved performance in Danish, with smaller spillovers to math exams. Importantly, the effects persist into early adulthood: diagnosed students are 3.0 percentage points more likely to complete an academic high school track before age 20 and 6.7 percentage points less likely to be not in education, employment, or training (NEET), corresponding to relative increases of 16 and 49 percent, respectively. These are intent-to-treat estimates and therefore likely understate the true effect of diagnosis and support, particularly since some undiagnosed students just above the threshold may still access similar support informally. Our findings suggest that timely identification and mandated support can substantially mitigate the educational disadvantages associated with dyslexia, with large effects that can meaningfully shift human capital accumulation over the life course.

Second, we find that diagnosis is initially associated with a decline in well-being spanning several measures, including school well-being and academic self-perception. The initial decline is approximately 0.09 SD for both measures. This decline dissipates within two years, and by four years after diagnosis, effects turn positive and statistically significant. While school well-being increases by approximately 0.09 SD, academic self-perception increases by approximately 0.18 SD, relative to marginally undiagnosed peers. This pattern is consistent with an initial period of adjustment to receiving a diagnosis that is outweighed, over

time, by the benefits of mandated support. The improvement in academic self-perception may be an important channel through which diagnosis and support translate into improved educational attainment, as growing confidence in one's own academic abilities may make students more willing to pursue more academically demanding education trajectories. Together, these results suggest that the well-being benefits of diagnosis and support take time to materialize, but ultimately exceed any short-run costs.

Third, we find evidence consistent with dynamic complementarities in the effects of diagnosis. Students identified earlier in compulsory schooling experience systematically larger and more persistent gains in educational outcomes than those identified later, consistent with early intervention increasing the returns to later skill formation (Cunha and Heckman, 2007; Heckman et al., Forthcoming). The same pattern extends to well-being, where students identified earlier have more time for the gains we document to accumulate before the end of compulsory schooling, suggesting that early identification may be a policy priority.

To the best of our knowledge, this paper is the first to identify the causal effect of receiving a dyslexia diagnosis using variation from an institutional classification threshold. We do so using population-level register data covering a full population of students, which allows us to trace effects from short-run reading performance through, most importantly, educational attainment and labor market status in early adulthood. By documenting substantial and persistent effects on educational attainment for a group whose potential is often constrained less by ability than by delayed or inconsistent identification (Snowling, Hulme and Nation, 2020), our findings imply potentially large long-run economic returns to timely diagnosis (Heckman, Humphries and Veramendi, 2018). Conversely, failing to identify and support dyslexic children risks leaving this potential untapped, implying losses for individual welfare and aggregate productivity.

More broadly, our findings contribute to an emerging literature examining the value of *marginal* diagnoses and classifications for individual outcomes (Almond, Lee and Schwartz, 2016; Iizuka et al., 2021; Kim, Lee and Lim, 2019; Alalouf, Miller and Wherry, 2024; Persson, Qiu and Rossin-Slater, 2025). This literature generally finds limited or null average effects of marginal diagnosis on the outcomes it is meant to improve. For instance, Almond, Lee and Schwartz (2016) find that classifying New York City students as overweight has little effect on subsequent weight outcomes, while Alalouf, Miller and Wherry (2024) find that a marginal type 2 diabetes diagnosis raises healthcare spending without producing lasting improvements in clinical measures. Persson, Qiu and Rossin-Slater (2025) similarly find, using variation from relative age within grade that shifts the likelihood of a marginal ADHD diagnosis, that cousins of marginally diagnosed children are no better off in the long run on educational and labor market outcomes, and possibly worse off. One interpretation of these null or negative findings is that clinical diagnostic thresholds are

set too low relative to where treatment is actually beneficial, rather than that diagnosis itself is uninformative. In contrast, we find that marginal diagnosis meaningfully improves both educational attainment and well-being. This divergence is consistent with the differing institutional role of diagnosis across settings. In much of this literature, diagnosis primarily conveys information or triggers a medical response contingent on further action by the patient, provider, or school, whereas in our setting a dyslexia diagnosis directly triggers legally mandated educational support.

Our findings also speak to a broader literature on how neurodevelopmental conditions shape children's long-run human capital trajectories. While dyslexia differs substantially from conditions such as attention-deficit/hyperactivity disorder (ADHD) and attention deficit disorder (ADD) in its underlying nature and treatment, existing evidence on these conditions documents substantial negative associations with human capital accumulation (Currie and Stabile, 2006) and adverse adult labor market outcomes (Fletcher, 2014). Relative to this literature, which has focused primarily on academic and labor market consequences, we additionally document effects on children's well-being, a dimension of diagnosis-related outcomes that remains largely unexplored across neurodevelopmental conditions more broadly.

Our findings also inform the literature on rules-based versus discretionary practices in diagnosis and treatment. Marquardt (Forthcoming) shows that physician-level discretion in setting ADHD diagnostic thresholds contributes to disparities in diagnosis by patient gender, and Cuddy and Currie (2026) document how discretion in adherence to treatment guidelines, once an adolescent has been diagnosed with a mental illness, shapes their subsequent outcomes in the United States. Unlike these settings, our diagnostic threshold is rule-based rather than left to individual judgment, which allows us to identify the causal effect of crossing into diagnosis using quasi-random variation at the cutoff. This includes the effect of any discretion exercised downstream in how support is implemented, since our estimates capture the total effect of the diagnosis-triggered bundle rather than requiring that its administration be uniform.

The remainder of the paper unfolds as follows. Section I describes the institutional setting, the dyslexia testing procedure, and the mandated support. Section II presents the data and sample restrictions. Section III details the empirical strategy. Section IV presents the main results. Section V discusses potential mechanisms, and Section VI concludes.

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 is associated with deficits 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 initial consonant sound illustrates phonological processing. This deficit impairs children's ability to decode 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 becomes more efficient and automatized over time, and eventually becomes effortless, dyslexic children may continue to struggle even with common, high-frequency words they have encountered repeatedly (Snowling, Hulme and Nation, 2020). Even for proficient readers with dyslexia, acquiring new vocabulary or learning an additional language may be more demanding relative to typically developing peers. These difficulties can make reading and writing slow and exhausting processes, leading to frustration and undermining academic self-concept.

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 of severity, rather than as a simple binary condition, making it challenging to establish unambiguous criteria for diagnosis based solely on reading ability (Gabrieli, 2009).[2] Historically, dyslexia has been defined as an unexpected, persistent, and specific learning disability that results in slow

[2]The diagnosis is further complicated by the fact that 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.

and/or inaccurate word recognition and spelling despite adequate instruction and intelligence (Lyon, Shaywitz 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 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 affects prevalence estimates (Wagner et al., 2020). Consistent 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 taken into account (Shaywitz, 1998). Based on this, there is no definitive prevalence estimate that can be considered universally correct.

Although dyslexia cannot be cured, many individuals develop compensatory strategies that partially offset their reading difficulties. Research suggests that increased reliance on right-hemisphere neural pathways, which are associated with visual processing and memory, may facilitate such compensation (Hoeft et al., 2011). Visualization strategies can support word recognition and comprehension by strengthening mental representations, while memory-based strategies may aid in learning specific word forms and spelling patterns. Early identification may therefore increase the returns to compensatory investments in skill formation by enabling children to adopt effective learning strategies at an earlier stage (Gabrieli, 2009).

B. Institutional Setting

We study children in compulsory school in Denmark. Compulsory schooling begins in August of the calendar year in which children turn 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. In 2021, 22 percent of the children who had their first day

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

of school entered a private school. Children are typically assigned to classes when entering preschool and usually remain 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 comparable to the OECD average. While the preschool class is organized around a grade teacher, Grades 1-9 are taught by subject-specific teachers that often follow the same cohort over multiple years.

Compulsory schooling ends with mandatory final examinations in seven subjects in Grade 9.[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 public 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, adjusting question difficulty to the student's performance. The well-being surveys are administered electronically, with a designated teacher responsible for conducting the survey in the class.[6] These assessments allow teachers to track the learning progress and well-being of each child and identify students who may need further assessment for special needs. Appendix A documents the timing of these assessments.[7] After compulsory schooling, students can enroll in either academic high school or vocational school. While an academic high school degree provides access to higher education, a vocational degree primarily qualifies for skilled jobs in the labor market.[8]

The Danish school system is based on a principle of inclusion, whereby students with special needs are educated within the general system whenever possible. This principle entails that the educational support in terms of special education is tailored to individual needs based on formal assessment. Specifically, the need for special education is determined through an evaluation by the municipal educational-psychological counseling service, which assesses both the presence and nature of the child's needs. The school principal, in consultation with parents, then decides whether the child should be referred to special education. If referred, the support is reassessed at least once a year.

During the period 2015–2019, 3.56 percent of students with dyslexia received formal special education. However, special education is also provided for a range of other condi-

[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 consist of subexams.

[6] 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.

[7] 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.

[8] A vocational degree can, under certain conditions, also provide access to specific higher education degrees that take practical approaches.

tions, including ADHD, ADD, intellectual disabilities, and hearing impairments. Students requiring more than nine hours of support per week typically receive instruction outside the regular classroom, either as a supplement to or as replacement for regular instruction in specific subjects.[9] In some cases, students receive all instruction in specialized classes, either within their school or at institutions dedicated to 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 for all children in compulsory schooling 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 phonological 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

[9]Students requiring fewer than nine hours of support typically remain in the classroom and receive assistance through differentiated instruction or the presence of a teaching assistant. Teaching assistants can be particularly beneficial in classrooms with behavioral challenges (Andersen et al., 2020*a*).

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 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 labeled as having inaccurate phonological coding suggests that you should get 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 and a complicated algorithm, 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.[10] If the test is not approved, the student can retake it when deemed appropriate.

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, 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), and children with limited proficiency in Danish as a non-native language. These limitations underscore the necessity for the sample selection as described in section II.

Alternative dyslexia tests exist apart from the standardized dyslexia test. As the standardized test is not supposed to be retaken, some children, parents, or teachers who are

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

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 standardized 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 standardized test if the standardized test has not proven the diagnosis. 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 principal 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.[11]

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 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 linked via a unique personal identifier assigned to every resident of Denmark. The data on dyslexia are from two distinct and previously unavailable sources. First, we obtained data from the Danish Ministry of Children

[11]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.

and Education on the standardized dyslexia test, containing all validated tests between 2015 and 2019. This data set includes the dyslexia scores and corresponding test dates. It also includes responses to a short pre-test questionnaire completed by students prior to taking the test. The questionnaire provides information, for example, on whether the student already receives additional educational support and whether the student believes he or she has dyslexia. Second, we gained access to the government-provided membership database for dyslexia-specific educational materials[12] that enables us to observe the enrollment of children with documented dyslexia.

We link these two data sources with administrative registers provided by Statistics Denmark, which we observe through the end of 2025. Among other things, the registers contain information on socioeconomic status and school-specific characteristics. They also include data on the standardized tests (available since 2010) and well-being surveys (available since 2015) conducted in public schools, as well as results from the school-leaving exams at the end of Grade 9. We further observe students' subsequent educational and labor market outcomes after compulsory schooling.

Leveraging these data, we construct primary outcome measures that capture short-, medium-, and long-run outcomes. For short-run effects, we examine performance on the first standardized reading test following diagnosis. Medium-run outcomes are measured using school-leaving exam grades at the end of Grade 9. Long-run outcomes capture educational and labor market attachment at age 20, including completion of an academic high school degree and whether the individual is not in education, employment, or training (NEET).[13] The latter serves as an indicator of youth disengagement, reflecting a failure to transition successfully from compulsory schooling into stable educational or employment pathways (Bynner and Parsons, 2002).

In addition to educational outcomes, we examine two main measures of well-being from the national well-being survey, which is conducted every year in Danish public schools for students in Grades 4 to 9. First, school well-being is a composite index of 29 survey questions grouped into four subscales: academic well-being, social well-being, support and inspiration in teaching, and calm and order in the classroom. The Ministry of Children and Education constructed these subscales using factor analysis, reliability analysis, and tests for differential item functioning (Ministry of Children and Education, 2016). Academic well-being covers students' perceptions of their academic skills, concentration, and problem-solving. Social well-being covers the sense of belonging to the school and class, feelings of safety, and experiences of bullying. Support and inspiration in teaching covers motivation, influence over what is taught, and perceptions of teachers' help and support.

[12]Known as NOTA in Danish.

[13]Given that our sample covers diagnoses between 2015 and 2019, long-run outcomes are observed for students diagnosed in earlier cohorts or at older ages, and the effective sample size is correspondingly smaller.

Calm and order covers noise in the classroom and classroom management. Appendix Table C2 lists the questions in each subscale, and Appendix Figure E1 shows results for each subscale separately, which broadly follow the same pattern as the composite index, with an initial decline followed by a gradual recovery. Second, academic self-perception captures students' assessment of their own academic competence, based on two questions: "I am doing well in school work" and "I am making good academic progress in school." These two questions are also part of academic well-being and hence of the composite. Academic self-perception is a strong predictor of pursuing further academically intensive education in Denmark, even conditional on exam performance (Gensowski et al., 2021).

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 again.

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 (i.e. between March and June) 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 the causal effects of a dyslexia diagnosis and the mandated support it triggers on educational outcomes and well-being for children at the margin of diagnosis. To do so, we exploit the scoring threshold of the standardized dyslexia test in a regression discontinuity (RD) design, comparing children who score just above and just below the age-specific diagnostic cutoff. The key identifying assumption is that all other determinants of outcomes vary continuously at the cutoff, which is plausible if children cannot precisely manipulate their score around it (Lee and Lemieux, 2010). Variation in diagnosis near the cutoff is then as good as random and fits the essence of a randomized experiment. As discussed in Section I.C, the test is computer-based and scored automatically based on a complicated algorithm, which makes precise manipulation of the

score difficult.

For each outcome y_{it}, we estimate the following model:

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

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 assumption that the distribution of ϵ_{it} is continuous at the cutoff, β is the intention-to-treat (ITT) effect in a sharp RD design of crossing the diagnostic cutoff, that is, of being diagnosed under the standardized testing system and receiving the support this triggers.

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 5.4 and 8.0 points, with a mean of 6.6, across the outcomes we study (see Appendix Table D1). Based on this, we use a common bandwidth of 7 points on either side of the cutoff for all regressions. 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 (sharp RD) coefficients are of interest in their own right, some children above the cutoff acquire a diagnosis through alternative testing methods, which introduces noncompliance and likely leads the reduced form to understate the effect of diagnosis and support. To address this, we complement the sharp RD with a fuzzy RD design, using enrollment in the government-provided database for dyslexia-specific materials as a proxy for realized diagnosis and support and crossing the cutoff as an instrument for enrollment.

The first stage captures the jump in enrollment induced by crossing the cutoff, estimated using the same functional form, kernel, and bandwidth as the reduced form. The ratio of the reduced-form coefficient to the first-stage coefficient identifies a local average treatment effect (LATE) for compliers, children whose enrollment status is induced to change by crossing the cutoff, under two further assumptions. Monotonicity requires that crossing

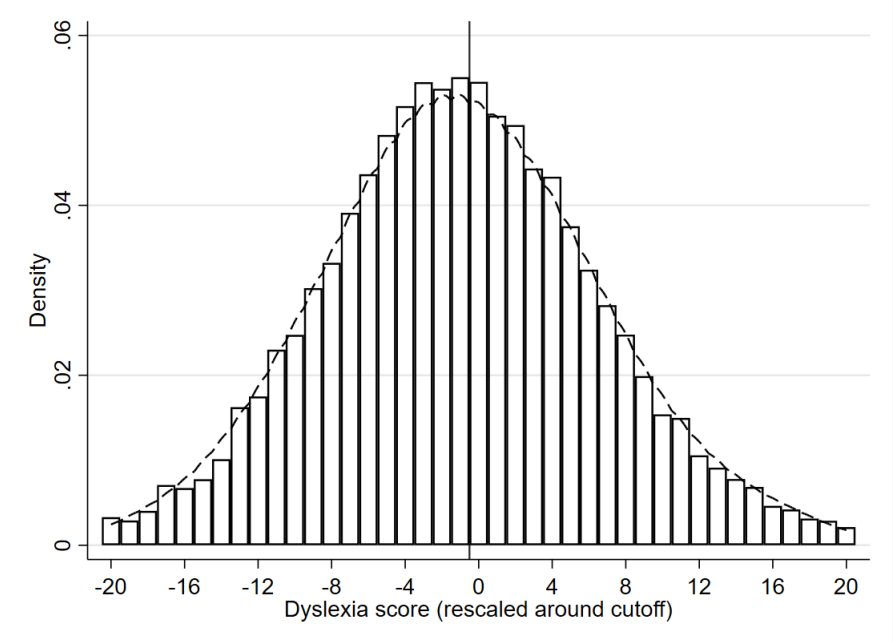


Figure 1. : Frequency of observations around the dyslexia cutoff

the cutoff does not make a child less likely to enroll. This assumption of no defiers seems likely to hold in our setting. Additionally, the exclusion restriction requires that crossing the cutoff affects outcomes only through diagnosis and the support it triggers, of which enrollment is one indicator. Because enrollment does not capture the diagnosis-triggered bundle in full, we interpret the LATE with some caution. In contrast, an advantage of the reduced-form model is that it requires fewer assumptions. In the absence of manipulation of the running variable and of other discontinuous changes at the cutoff, which we assess in Section III.A, it consistently captures the intent-to-treat effect for marginally diagnosed children.

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

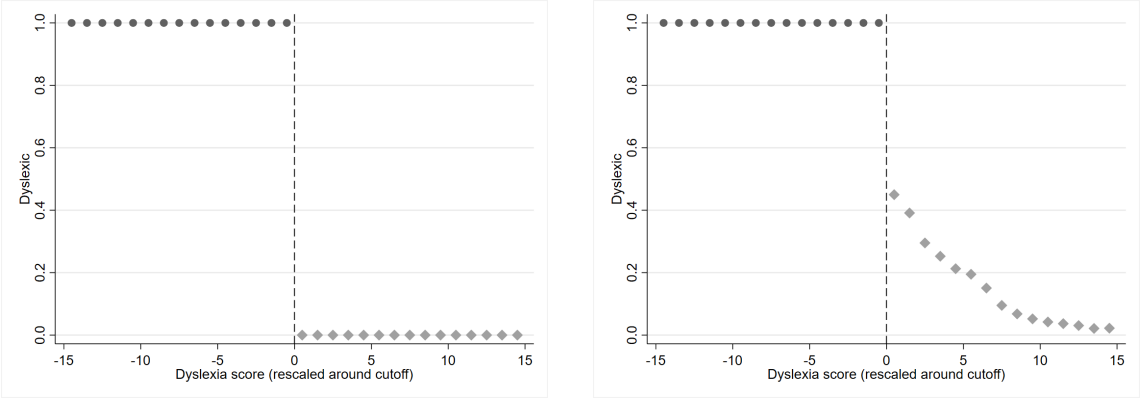
this statement.

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 dyslexic, 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 can generally verify this condition. Of the predetermined covariates we examine, two show a statistically significant discontinuity at the 5 percent level: children just below the cutoff are less likely to live with both parents and have a lower standardized math score than children just above it. Four further covariates show a marginally significant discontinuity at the 10 percent level: self-reported difficulty reading, academic self-perception, father's employment, and father's income. Given the number of covariates tested, this pattern is broadly consistent with what would be expected under random chance at conventional significance levels. The discontinuity in math score plausibly reflects that the running variable, which measures phonological reading ability, is correlated with general academic performance, rather than a threat to identification specific to our outcomes of interest. The discontinuity in family structure is harder to attribute to a specific mechanism, and we return to this in our sensitivity analysis, where we show that our main results are robust to the inclusion or exclusion of these covariates. For the remaining covariates, there is no evidence of a discontinuity around the cutoff at the 5 percent significance level.

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

[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.

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 the main empirical findings. We first report first-stage estimates that document how crossing the diagnostic cutoff translates into receiving dyslexia support. We then present the effects of diagnosis on academic performance and well-being, followed by a series of robustness checks.

A. First Stage

The Danish dyslexia test generates a sharp discontinuity in diagnosis: students scoring below the cutoff are classified as dyslexic, while those above are not, as illustrated in Figure 2a. Because some children obtain a diagnosis outside the standardized testing system, we use this discontinuity to scale our reduced-form estimates in a fuzzy RD framework, where treatment is defined as enrollment in the government-provided database for dyslexia-specific materials, our proxy for realized diagnosis and support.

Panel (b) of Figure 2 plots the probability of receiving dyslexia support against the test

score. Children to the left of the cutoff are diagnosed through the standardized test and are automatically assigned support. However, as detailed in Section I.C, some students above the cutoff obtain a diagnosis through alternative assessments and may therefore also receive support, which explains the positive probability of treatment to the right of the cutoff. The probability of receiving support declines smoothly as the test score increases above the cutoff, consistent with students near the threshold exhibiting similar underlying reading difficulties. Due to the test's random nature, a child scoring just above the cutoff on a given test day may well have scored below it on another, and may continue to show symptoms of dyslexia in daily life despite not receiving a formal diagnosis. This generates imperfect compliance with the diagnostic rule, which we instrument for with an indicator for crossing the cutoff.

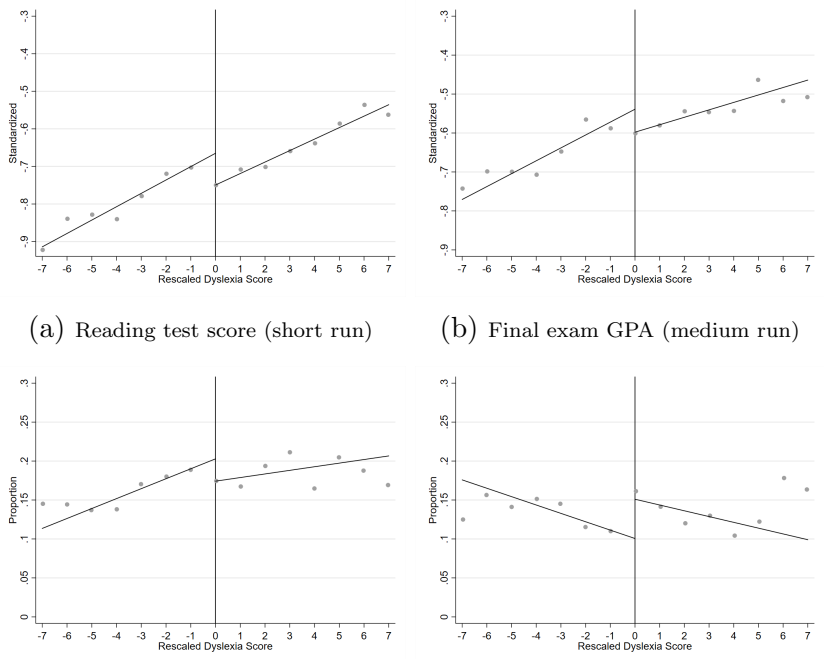
Column (1) of Table 1 reports the first-stage estimates. Crossing the diagnostic cutoff increases the probability of receiving dyslexia support by approximately 53 percentage points. The estimates are precisely estimated and statistically significant, indicating a strong first stage and alleviating concerns about weak instruments in the fuzzy RD design.

B. Educational outcomes

Figure 3 provides visual evidence on the relationship between dyslexia diagnosis and educational outcomes across the full schooling trajectory. Each panel shows first-degree polynomial regression lines estimated separately on each side of the cutoff, with dots representing conditional means within one-unit bins.

Panel (a) focuses on the short-run outcome, showing a clear discontinuity in reading test scores following diagnosis. Panel (b) turns to the medium run, where diagnosed students achieve higher GPA at the end of compulsory schooling. Panel (c) documents the long-run effect on academic high school completion, while panel (d) shows a reduction in NEET status at age 20. Across all four panels, the discontinuity at the cutoff is visually apparent, with the regression lines shifting upward (or downward for NEET) to the left of the threshold.

Column (2) of Table 1 presents the corresponding reduced-form regression results from our baseline model. Each estimate reports the coefficient from a separate local-linear regression with a triangular kernel, a bandwidth of 7 points, and covariates included. We report conventional and bias-corrected estimates, robust standard errors, the mean of the outcome for observations above the cutoff, and the effective number of observations (N). Consistent with the graphical evidence, we find statistically significant coefficients across all four outcomes. In the short run, diagnosed students achieve higher reading performance, increasing test scores by 0.097 SD. In the medium run, these gains extend to school-leaving exam GPA, which rises by 0.063 SD. The effects persist into early adulthood: diagnosed



(a) Reading test score (short run) (b) Final exam GPA (medium run)

(c) Academic HS completion (long run) (d) NEET at age 20 (long run)

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

Notes: The figure displays separate first-degree polynomial estimates on each side of the dyslexia cutoff, based on a bandwidth of 7 points. Each dot represents the average outcome within a one-unit bin. The x-axis shows the dyslexia test score rescaled around the cutoff, with negative values indicating a dyslexia diagnosis.

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

	First stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean (4)	N (5)
Panel A: Short run					
Reading score	**0.524**	**0.094**	**0.176**		
	[0.532]	**[0.097]**	**[0.184]**	-0.657	15,142
	(0.006)	(0.017)	(0.032)		
Panel B: Medium run					
GPA	**0.540**	**0.055**	**0.101**		
	[0.547]	**[0.063]**	**[0.117]**	-0.544	16,638
	(0.006)	(0.015)	(0.027)		
Panel C: Long run					
High school completion	**0.507**	**0.023**	**0.045**		
	[0.513]	**[0.030]**	**[0.058]**	0.188	12,354
	(0.009)	(0.007)	(0.014)		
NEET at age 20	**0.497**	**-0.053**	**-0.105**		
	[0.503]	**[-0.067]**	**[-0.135]**	0.138	6,257
	(0.011)	(0.008)	(0.016)		

Sample of children with dyslexia score within a 7 point average around the cutoff. Each cell reports the coefficient from a separate local-linear regression with a triangular kernel. Bias-corrected estimates are listed in square brackets and robust standard errors in parentheses below. The mean of the outcome is reported for children with a dyslexia score above the cutoff. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).

students are 3.0 percentage points, or 16 percent relative to the mean, more likely to complete an academic high school track before age 20, and 6.7 percentage points, or 49 percent relative to the mean, less likely to be NEET at age 20.

Appendix Table E1 disaggregates the GPA effect by exam subject. The improvement is driven primarily by Danish (0.081 SD) and math (0.065 SD), both significant, while we find no corresponding improvement in English (-0.007 SD), where dyslexic students receive less of the mandated support available in other subjects. As the dyslexia support targets skills in reading, writing, and spelling, this pattern suggests a direct effect of the support, concentrated in the subjects where these skills are most directly assessed, rather than a generic improvement across all subjects.

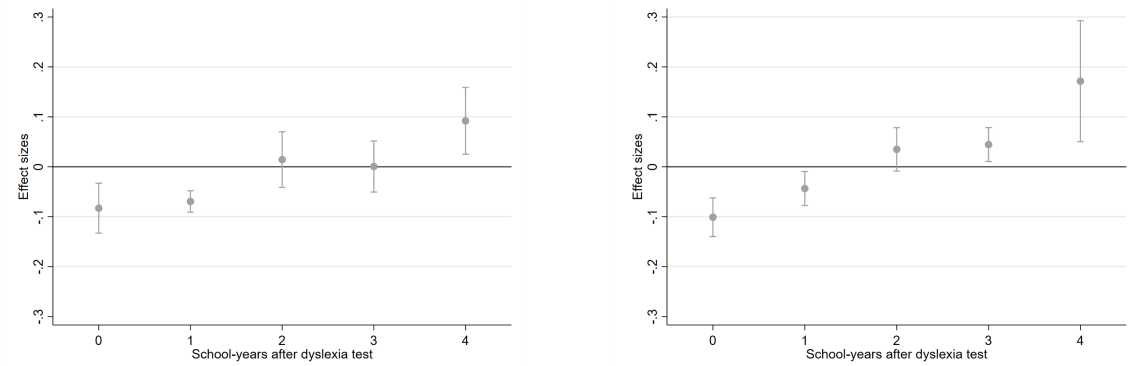
The Danish and math exams consist of multiple subexams. Appendix Table E2 shows the regression results for each of the subexams. While the other subexams are mandatory, the oral exam in math is one of the subjects randomly drawn from a pool of possible exam subjects, so the sample size around the cutoff is substantially smaller than for the mandatory exams. We observe statistically significant effects on the written, spelling, and reading subexams in Danish, and on the basic skills and problem-solving subexams

in math, but no significant effect on either the Danish or the math oral subexam. This pattern is consistent with the effect operating primarily through written, reading-based skills, the domain most directly targeted by dyslexia-specific support, rather than through oral communication skills.

In column (3), we divide the reduced-form coefficient in column (2) by the first-stage coefficient in column (1) to obtain LATE estimates. In all cases where the reduced-form estimates are statistically significant, the same holds true for the LATE estimates, which are mechanically larger. These are intent-to-treat estimates and therefore likely understate the true effect of diagnosis and support, particularly since some undiagnosed students just above the threshold may still access similar support informally.

C. Well-being outcomes

Next, we analyze the effect of diagnosing dyslexia and providing mandated support on well-being. Specifically, we consider two main measures: overall school well-being and academic self-perception. Figure 4 shows the two measures in the year of diagnosis and the following four years.



(a) School well-being (b) Academic self-perception

Figure 4. : Well-being

The figure presents bias-corrected reduced-form estimates along with robust confidence intervals. The estimates are based on a sample of children with dyslexia scores within a 7-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.

Diagnosis is initially associated with a decline in both measures, of approximately 0.09 SD. This likely reflects a period of adjustment following diagnosis, consistent with children initially processing a new label for their difficulties before benefiting from the support it triggers. The decline dissipates within two years, and by four years after diagnosis,

both measures turn positive and statistically significant relative to marginally undiagnosed peers: school well-being increases by approximately 0.09 SD, while academic self-perception increases by approximately 0.18 SD. For academic self-perception, this pattern is consistent with growing confidence in one's own academic abilities as students become more comfortable with the tools and support available to them and experience improved academic performance over time. The improvement in academic self-perception may be an important channel through which diagnosis and support translate into improved educational attainment, as growing confidence in one's own abilities may make students more willing to pursue more academically demanding education trajectories.

Appendix Figure E1 decomposes school well-being into its four underlying subscales: academic well-being, social well-being, support and inspiration in teaching, and calm and order in the classroom. All four follow the same broad pattern as the composite, with an initial decline, a return to zero within two years, and a positive estimate four years after the test. The initial decline is of similar size (roughly 0.07 to 0.09 SD) for academic well-being, support and inspiration, and calm and order, and for the latter two it persists into the following year. For social well-being, the decline is delayed: the estimate is small and insignificant in the year of the test and reaches roughly 0.10 SD one year later. Four years after the test, the estimates are positive and statistically significant for academic well-being (approximately 0.16 SD), social well-being (approximately 0.09 SD), and calm and order (approximately 0.08 SD). The estimate for support and inspiration is also positive (approximately 0.05 SD) but not statistically distinguishable from zero. The recovery is thus broad-based, but the largest gain is in academic well-being, which is consistent with the pattern for academic self-perception, since the two questions that make up that measure are part of this subscale.

D. Sensitivity checks

Appendix Table F1 reports reduced-form estimates for our four main outcomes under alternative specifications, and Appendix Figure F1 shows how the estimates vary with bandwidth and polynomial order.

Covariates. — Our main specification includes predetermined covariates. If the RD assumptions hold, omitting them should leave the point estimates unchanged and affect only precision. Column (2) shows that this is the case: the estimates change by at most 0.009 (0.001 for NEET) and all remain significant. Precision is similar with and without covariates, except for GPA, where the standard error roughly doubles (from 0.015 to 0.029).

Bandwidth and functional form. — In column (3), we let the bandwidth differ across outcomes using the optimal bandwidth of Calonico, Cattaneo and Titiunik (2014), which ranges from 5.4 to 8.0 points (Appendix Table D1). The estimates are somewhat larger for all outcomes, most notably for NEET (-0.070 compared with -0.053), and all remain significant. In column (4), we use a second-degree polynomial in the test score. The estimates are again larger for all four outcomes, for instance 0.129 for reading and 0.039 for high school completion, and all remain significant. Appendix Figure F1 shows the bias-corrected estimates for bandwidths from 5 to 15 points and both polynomial orders. The reading and GPA estimates are stable across the full range. The estimates for high school completion decline modestly as the bandwidth widens, and a few become marginally insignificant at the widest bandwidths. For NEET, the linear estimates shrink as the bandwidth grows, from about -0.09 at 5 points to about -0.03 at 15 points, and lose significance only at bandwidths of roughly 13 points or more, well beyond the data-driven optimum. The second-degree estimates for NEET remain between about -0.06 and -0.09 and significant throughout, which suggests that the attenuation of the linear estimates at wide bandwidths reflects curvature in the outcome rather than a fragile effect at the cutoff.

Heaping. — In column (5), we estimate 'doughnut' regressions that exclude children scoring exactly at the cutoff, which addresses the concern that excess mass at particular values of the running variable could bias the estimates if it is not symmetric around the threshold. The estimates for reading, GPA, and high school completion are virtually unchanged. The NEET estimate is smaller (-0.038 compared with -0.053) and less precise, but remains significant.

Kernel. — Our main specification uses a triangular kernel. In column (6), we use a rectangular kernel, which gives equal weight to all observations within the bandwidth. The estimates for reading, GPA, and NEET remain significant and of the same sign, but those for GPA (0.039) and NEET (-0.023) are smaller than in the baseline. The estimate for high school completion falls to 0.008 and is no longer statistically significant. The high school result is thus the least robust of our main findings to the choice of kernel, and we place more weight on the triangular kernel, which gives the greatest weight to observations close to the cutoff.

Summary. — The estimates for reading, GPA, and NEET are significant in all six specifications, and the estimate for high school completion is significant in five of six. The magnitudes vary across specifications, but the signs are stable, and the alternative bandwidth and polynomial choices tend to produce larger rather than smaller effects.

V. Suggestive Mechanisms

Our main results show that diagnosis and mandated support raise reading performance, exam outcomes, and educational attainment, and that an initial decline in well-being reverses within two years. We now examine whether these effects differ across groups of children. All estimates in this section are reduced-form RD estimates in subsamples.

We first ask whether the timing of diagnosis matters. Appendix Table G1 splits the sample by the grade in which the child was diagnosed: elementary school (Grades 3 to 6) and lower secondary education (Grades 7 to 8). The effect on the next reading test is 0.097 SD and significant for elementary school students, compared with 0.043 SD and insignificant for lower secondary students. The GPA effect points in the opposite direction, at 0.050 SD for elementary school and 0.073 SD for lower secondary, and is significant only for the former, which reflects a difference in precision rather than in magnitude. The GPA results therefore give no evidence that earlier diagnosis produces larger gains. For the long-run outcomes, the effect on completing an academic high school track is 3.2 percentage points for students diagnosed in elementary school and essentially zero (0.2 percentage points) for those diagnosed in Grades 7 to 8. NEET at age 20 falls significantly in both groups, by 7.3 percentage points for elementary school and 3.4 percentage points for lower secondary. Overall, earlier diagnosis is associated with larger effects on reading, high school completion, and NEET, but not on GPA. Timing also matters for well-being. The gains in school well-being and academic self-perception emerge only about four years after the test (Figure 4), and the well-being survey covers Grades 4 to 9, so these gains can be observed only for students diagnosed by around Grade 5. Students diagnosed earlier therefore have more time for the gains to accumulate before the end of compulsory schooling, whereas for later-diagnosed students they would arrive, if at all, after the survey window closes.

Appendix Table G2 reports estimates by gender and by SES, where low and high SES are defined relative to the sample median of a combined index of parental education and income rank. The reading and GPA effects are significant in all four subgroups except for low-SES reading, and the differences between groups are small: 0.085 SD for boys and 0.105 SD for girls in reading, and 0.062 SD and 0.047 SD in GPA. The effect on high school completion is nearly identical for boys and girls (2.1 and 2.3 percentage points, the latter only significant at the 10 percent level), while the reduction in NEET is larger for girls (7.4 percentage points, compared with 3.6 for boys). By SES, the reading estimate is smaller for low-SES students (0.076 SD, compared with 0.105 SD) and not statistically significant, but it is also less precisely estimated. The GPA effect is identical (0.054 SD in both groups). In the long run, the point estimates are roughly twice as large for high-SES students: 3.2 compared with 1.6 percentage points for high school completion, and a reduction in NEET of 6.6 compared with 3.1 percentage points. All four long-run

estimates are significant, so low-SES students clearly benefit, but the gains in attainment do not appear to be concentrated among them. This suggests that diagnosis and support alone are unlikely to close SES gaps in attainment, although we cannot rule out equal effects given the overlapping confidence intervals.

We next examine whether the effects carry over to the marks that subject teachers assign before the final exams (Appendix Table G3). Unlike the written exams, these marks are not set or graded centrally, and teachers know which students are diagnosed. The pattern mirrors the exam results for Danish (0.057 SD) and math (0.048 SD), which are both significant, and for English, where we again find no effect (0.013 SD). The estimate for science is small (0.018 SD) but significant, whereas the corresponding exam estimate is not (Appendix Table E1). The average across subjects is smaller than the exam GPA effect (0.019 SD compared with 0.055 SD) and significant only at the 10 percent level. The agreement between teacher marks and centrally set exams for Danish and math suggests that the exam results do not merely reflect differences in the accommodations available at the exam.

Finally, Appendix Table G4 examines whether the estimated effects could reflect changes in schooling arrangements after diagnosis. The estimates for enrollment in special education are small and not significant at the 5 percent level in any of the four years. The estimates for private schooling are small and do not follow a consistent direction: slightly negative in year 0 (-0.4 percentage points), positive in year 1 (0.6 percentage points), insignificant in year 2, and negative in year 3 (-1.1 percentage points), against a base rate of 13 to 15 percent. We therefore see no evidence that diagnosis systematically shifts students into special education or between school types, and such shifts are unlikely to explain our main results.

VI. Conclusion

Many children leave compulsory school without basic reading skills, and for some of them the cause is dyslexia, a condition that is common but hard to diagnose consistently. Evidence on what a diagnosis does for these children has been scarce, largely because diagnoses in most settings depend on subjective assessments and on selection into testing. We exploit the nationwide standardized dyslexia test introduced in Denmark in 2015, which assigns diagnosis and legally mandated support according to a rule-based threshold, and use a regression discontinuity design to estimate the effects of diagnosis and support for children at the margin.

Our findings contribute to the literature on human capital formation in three ways. First, diagnosis and mandated support generate persistent improvements in educational outcomes. Marginally diagnosed students score 0.097 SD higher on the next reading test

and 0.063 SD higher in the school-leaving exam GPA. The effects extend into early adulthood: diagnosed students are 3.0 percentage points (16 percent) more likely to complete an academic high school track and 6.7 percentage points (49 percent) less likely to be NEET at age 20. These estimates are intent-to-treat and likely understate the effect of diagnosis and support, since some students just above the threshold obtain support through other channels. The results are robust across most specifications, although the high school completion estimate is not significant with a rectangular kernel, and the NEET estimate from a linear specification attenuates at bandwidths well beyond the optimum.

Second, we find evidence consistent with dynamic complementarities. Earlier diagnosis is associated with larger effects on longer run outcomes, in line with early intervention raising the returns to later investments in skill formation. Third, we provide new evidence on the well-being consequences of a diagnosis. Well-being declines by about 0.09 SD in the year of the test, and this decline dissipates within two years. Academic self-perception then improves, with a small positive estimate three years after the test and an estimate of about 0.18 SD after four years, and school well-being is higher after four years as well.

These findings have implications for the design of diagnosis and support in schools. Standardized, rule-based testing removes dependence on the assessor and the school, and in our setting a diagnosis that triggers mandated support produced lasting gains even for children who only just qualified. This contrasts with much of the literature on marginal diagnoses, where limited effects are found when diagnosis conveys information without directly triggering a defined response. It also suggests that children who score just above the threshold may face similar difficulties and could benefit from comparable support, which would be consistent with a threshold set too strictly. Our design cannot establish this, since it identifies effects only at the margin and cannot separate the diagnosis label from the support it triggers. Even so, our results indicate that timely diagnosis paired with mandated support can raise attainment and, after an initial adjustment period, improve how children experience school.

REFERENCES

Alalouf, Mattan, Sarah Miller, and Laura R Wherry. 2024. "What difference does a diagnosis make? Evidence from marginal patients." *American Journal of Health Economics*, 10(1): 97–131.

Almond, Douglas, Ajin Lee, and Amy Ellen Schwartz. 2016. "Impacts of classifying New York City students as overweight." *Proceedings of the National Academy of Sciences*, 113(13): 3488–3491.

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.

Bynner, John, and Samantha Parsons. 2002. "Social exclusion and the transition from school to work: The case of young people not in education, employment, or training (NEET)." *Journal of vocational behavior*, 60(2): 289–309.

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.

Chan, David C, Matthew Gentzkow, and Chuan Yu. 2022. "Selection with Variation in Diagnostic Skill: Evidence from Radiologists." *The Quarterly Journal of Economics*, 137(2): 729–783.

Cuddy, Emily, and Janet Currie. 2026. "Rules versus Discretion: Treatment of Mental Illness in US Adolescents." *Journal of Political Economy*, 134(1): 478–522.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review*, 97(2): 31–47.

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.

Daniel, Johny, Julian Elliott, Peter Tymms, and Steve Strand. 2026. "Specific Learning Difficulties: Disparities in Identification." *Journal of Learning Disabilities*, 59(5): 286–299. PMID: 41742682.

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.

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.

Gensowski, Miriam, Steven G. Ludeke, Oliver P. John, and Simon Calmar Andersen. 2021. "Academic self perceptions in a national Danish sample: Predictive power and development from grade 4 to 9." *Journal of Research in Personality*, 92: 104090.

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., and Stefano Mosso. 2014. "The Economics of Human Development and Social Mobility." *Annual Review of Economics*, 6(Volume 6, 2014): 689–733.

Heckman, James J., Haihan Tian, Zijian Zhang, and Jin Zhou. Forthcoming. "Dynamic Complementarity." *American Economic Journal: Applied Economics*.

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.

Houmark, Mikkel Aagaard, Victor Ronda, Esben Agerbo, Preben Bo Mortensen, and Michael Rosholm. 2025. "Genetic and Socioeconomic Achievement Gaps in Elementary School." *The Economic Journal*, 135(672): 2426–2454.

Iizuka, Toshiaki, Katsuhiko Nishiyama, Brian Chen, and Karen Eggleston. 2021. "False alarm? Estimating the marginal value of health signals." *Journal of Public Economics*, 195: 104368.

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

Kim, Hyuncheol Bryant, Suejin A. Lee, and Wilfredo Lim. 2019. "Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea." *Journal of Health Economics*, 65: 1–14.

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.

Marquardt, Kelli. Forthcoming. "Mis(sed) Diagnosis: Physician Decision-Making and ADHD." *Journal of Political Economy*, 0(0): 000–000.

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.

Ministry of Children and Education. 2016. "Metodenotat: Beregning af indikatorer i den nationale trivselsmåling i folkeskolen." Ministry of Children and Education, Copenhagen. Method note on the calculation of indicators in the national well-being survey in Danish public schools. Accessed September 2026.

OECD. 2026. "PISA 2025 Results (Volume I): Future-Ready Students." OECD Publishing, Paris.

Persson, Petra, Xinyao Qiu, and Maya Rossin-Slater. 2025. "Family Spillover Effects of Marginal Diagnoses: The Case of ADHD." *American Economic Journal: Applied Economics*, 17(2): 225–56.

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.

Yang, Liping, Chunbo Li, Xiumei Li, Manman Zhai, Qingqing An, You Zhang, Jing Zhao, and Xuchu Weng. 2022. "Prevalence of developmental dyslexia in primary school children: A systematic review and meta-analysis." *Brain sciences*, 12(2): 240.

Tests and Surveys

Table A1 summarizes the frequency of national tests and well-being surveys administered during compulsory schooling in Denmark. Reading tests have been administered annually since 2010 to students in Grades 2, 4, 6, and 8. Math tests are administered to students in Grades 3 and 6, and, since 2018, also in Grade 8. English tests are administered to students in Grade 7, and science tests to students in Grade 8 during the spring term. The well-being survey has been administered annually to students in Grades 4 through 9 since 2015.

Table A1—: Frequency of Tests and Surveys

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

Notes: The table shows the frequency of national tests and well-being surveys in Danish compulsory schooling.
[a] Grade 8 math testing was introduced in 2018.

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.

Data

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

Table C2—: Well-being survey questions by subscale

Question	Text
Academic well-being (8 questions)	
Q3	What do your teachers think of your progress in school?
Q5	Do you succeed in learning what you want to learn in school?
Q11	How often can you find a solution to problems if you try hard enough?
Q12	How often can you accomplish what you set out to do?
Q13	Can you concentrate during lessons?
Q14†	I am doing well in school work.
Q15†	I am making good academic progress in school.
Q29	If I am disturbed during class, I can quickly concentrate again.
Social well-being (10 questions)	
Q1	Are you happy at your school?
Q2	Are you happy in your class?
Q6	Do you feel lonely?
Q7	I feel that I belong at my school.
Q8	I like the breaks at school.
Q20	Most of the students in my class are friendly and helpful.
Q21	Other students accept me as I am.
Q23	Have you been bullied this school year?
Q25	Are you afraid of being laughed at in school?
Q26	How often do you feel safe at school?
Support and inspiration in teaching (7 questions)	
Q4	Do your teachers help you learn in ways that work well?
Q19	The teaching makes me want to learn more.
Q22	The teachers are good at supporting and helping me in school.
Q27	Do you and your classmates help decide what to work on in class?
Q28	The teachers make sure students' ideas are used in teaching.
Q31	Is the teaching boring?
Q32	Is the teaching exciting?
Calm and order (4 questions)	
Q30	If there is noise in the class, the teachers can quickly restore calm.
Q35	Do your teachers arrive on time for lessons?
Q36	Is it easy to hear what the teacher says in class?
Q37	Is it easy to hear what the other students say in class?

Notes: The table lists the 29 questions from the Danish national well-being survey (Grades 4 to 9) that form the Ministry of Children and Education's indicators, translated by the authors. School well-being is the simple average of all 29 questions, and each subscale is the simple average of its questions. A score is only computed if the student answered at least half of the questions. † Q14 and Q15 constitute academic self-perception. Source: Ministry of Children and Education (2016).

RD DESIGN

Table D1—: Bandwidth Selection

	Bandwidth estimate (1)	Bandwidth bias (2)
Short run		
- Reading score (std.)	5.859	10.395
Medium run		
- GPA, end of compulsory schooling (std.)	8.042	12.215
Long run		
- Completed high school before age 20	6.970	12.959
- Not in employment or education at age 20	5.394	10.628

Each cell indicates the bandwidth that minimizes the mean squared error between the predictions of two models regressing the outcome indicated in the row on the rescaled dyslexia test score (running variable). Column 1 reports the optimal bandwidth for point estimators, and column 2 reports the optimal bandwidth for inference.

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 dyslexia test					
- Difficult to learn reading and spelling? (1/0)	0.012	[0.011]	(0.009)	0.771	19,131
- Receive additional help in reading? (1/0)	-0.001	[-0.008]	(0.019)	0.376	19,131
- Previously received additional help? (1/0)	0.001	[-0.008]	(0.024)	0.608	19,131
- Do you have difficulty reading? (1/0)	*0.022*	*[0.020]*	(0.010)	0.610	19,131
- Do you have difficulty spelling? (1/0)	-0.002	[-0.008]	(0.008)	0.789	19,131
- Do you think that you are dyslexic? (1/0)	0.004	[0.001]	(0.011)	0.529	19,131
- Teacher told you that you are dyslexic? (1/0)	0.001	[0.002]	(0.009)	0.059	19,131
- Do you have tools that your peers don't? (1/0)	-0.016	[-0.026]	(0.018)	0.299	19,131
B. Child characteristics					
- Boy (1/0)	-0.022	[-0.022]	(0.014)	0.565	19,131
- Western ethnicity (1/0)	0.001	[0.000]	(0.005)	0.011	19,131
- Non-western ethnicity (1/0)	0.007	[0.009]	(0.008)	0.052	19,131
- Living with both parents (1/0)	**-0.023**	**[-0.030]**	(0.007)	0.698	19,131
- School starting age (years)	-0.008	[-0.008]	(0.005)	6.139	18,594
- Private schooling (1/0)	-0.004	[-0.004]	(0.004)	0.127	19,131
- Special needs teaching (1/0)	-0.001	[-0.001]	(0.001)	0.002	19,002
- Special needs classroom (1/0)	0.000	[-0.000]	(0.001)	0.001	19,002
- Psychiatric diagnosis (1/0)	0.000	[-0.000]	(0.002)	0.014	19,131
- Reading score (std.)	-0.006	[-0.015]	(0.031)	-0.678	16,401
- Math score (std.)	**-0.036**	**[-0.039]**	(0.014)	-0.419	14,972
- Conscientiousness (std.)	0.018	[0.018]	(0.066)	-0.210	6,413
- Academic self-perception (std.)	*-0.050*	*[-0.055]*	(0.030)	-0.257	5,710
- Happy to attend your school (std.)	-0.002	[0.014]	(0.020)	0.006	6,599
- Emotional stability (std.)	-0.050	[-0.049]	(0.040)	-0.042	6,264
- Peer relationship (std.)	-0.041	[-0.041]	(0.030)	-0.054	5,978
- Calm and order (std.)	-0.028	[-0.026]	(0.034)	0.057	6,207
- Support and inspiration (std.)	-0.020	[-0.017]	(0.025)	0.052	6,644
- Absence (percent)	0.003	[0.019]	(0.046)	1.612	16,619
- Age at dyslexia test date (years)	-0.053	[-0.038]	(0.050)	11.808	19,131
C. Mother's characteristics					
- Age at birth (years)	0.081	[0.121]	(0.101)	30.461	19,093
- Employed (1/0)	-0.006	[-0.007]	(0.009)	0.749	19,088
- High school or less (1/0)	-0.001	[0.003]	(0.008)	0.198	19,131
- Education length (years)	0.056	[0.037]	(0.052)	14.390	18,886
- Work experience (years)	0.024	[0.077]	(0.175)	14.369	18,997
- Income (1,000 DKK)	2.897	[4.923]	(4.621)	292.454	19,031
D. Father's characteristics					
- Age at birth (years)	0.018	[0.023]	(0.108)	32.999	18,890
- Employed (1/0)	*-0.013*	*[-0.017]*	(0.010)	0.852	18,910
- High school or less (1/0)	0.003	[0.009]	(0.011)	0.244	19,131
- Education length (years)	-0.031	[-0.058]	(0.053)	14.153	18,424
- Work experience (years)	-0.223	[-0.308]	(0.466)	18.748	18,566
- Income (1,000 DKK)	*-13.779*	*[-17.018]*	(10.119)	419.987	18,705

Sample of children with dyslexia score within a 7 point average around the cutoff. Columns 1-5 report reduced-form results from separate local-linear regressions similar to the main equation with the 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 variable 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 (Calonico et al., 2014; 2019).

Results

Table E1—: Effects on GPA and Exam Subjects, End of Compulsory Schooling

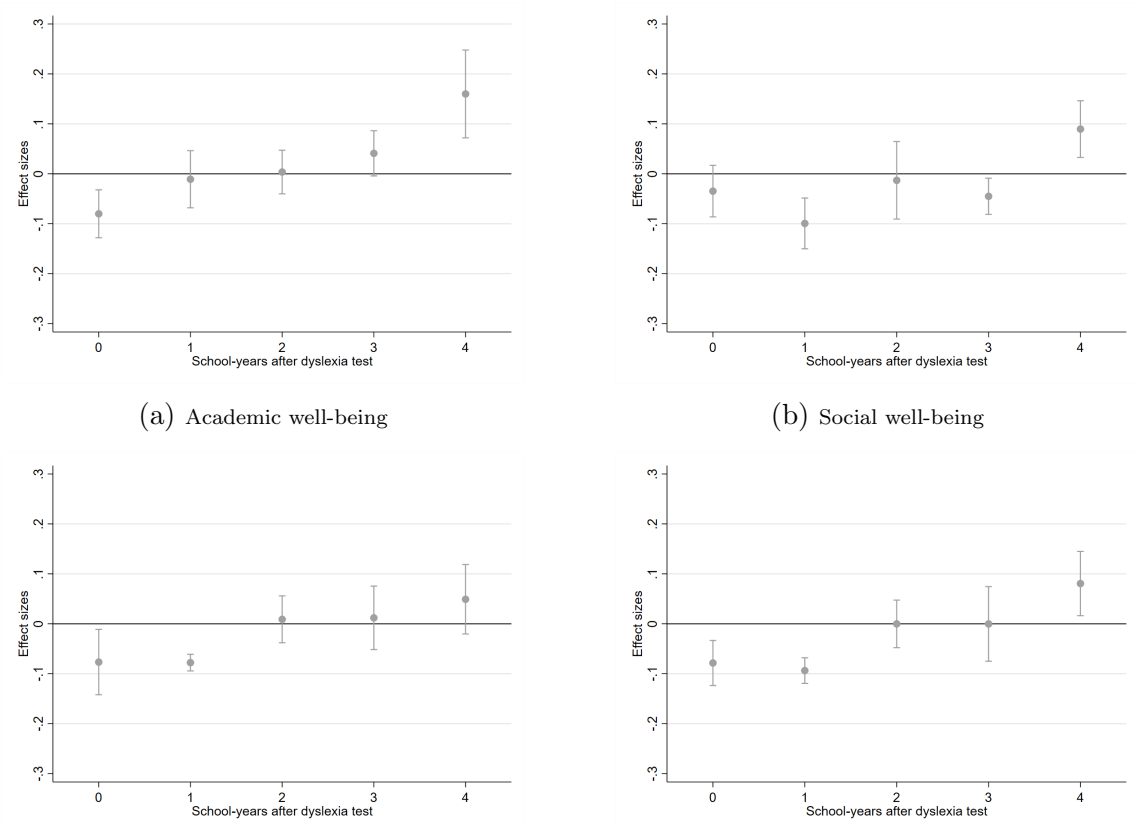
	First stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean (4)	N (5)
GPA	**0.540**	**0.055**	**0.101**		
	[0.547]	**[0.063]**	**[0.117]**	-0.544	16,638
	(0.006)	(0.015)	(0.027)		
Danish exam (std.)	**0.542**	**0.079**	**0.144**		
	[0.548]	**[0.081]**	**[0.149]**	-0.642	17,241
	(0.006)	(0.018)	(0.032)		
Math exam (std.)	**0.542**	**0.060**	**0.110**		
	[0.548]	**[0.065]**	**[0.121]**	-0.412	17,215
	(0.006)	(0.017)	(0.031)		
English exam (std.)	0.540	-0.011	-0.021		
	[0.547]	[-0.007]	[-0.013]	-0.540	16,903
	(0.007)	(0.014)	(0.026)		
Science exam (std.)	0.542	0.012	0.022		
	[0.548]	[0.026]	[0.047]	-0.325	16,637
	(0.006)	(0.021)	(0.038)		

Sample of children with dyslexia score within a 7 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 parentheses below. Mean of the outcome is reported for children with a dyslexia score above the cutoff. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).

Table E2—: Subexams for the Final Exams

	First stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean (4)	N (5)
Danish					
Written	**0.542**	**0.059**	**0.107**		
	[0.549]	**[0.071]**	**[0.130]**	-0.501	17,138
	(0.006)	(0.016)	(0.029)		
Spelling	**0.542**	**0.099**	**0.182**		
	[0.548]	**[0.098]**	**[0.182]**	-0.738	17,197
	(0.006)	(0.021)	(0.038)		
Oral	0.542	-0.003	-0.005		
	[0.548]	[-0.007]	[-0.012]	-0.391	17,107
	(0.006)	(0.019)	(0.035)		
Reading	**0.542**	**0.122**	**0.222**		
	[0.549]	**[0.125]**	**[0.231]**	-0.563	17,202
	(0.006)	(0.011)	(0.020)		
Math					
Basic skills	**0.542**	**0.065**	**0.118**		
	[0.548]	**[0.067]**	**[0.124]**	-0.425	17,203
	(0.006)	(0.024)	(0.044)		
Problem-solving	**0.543**	**0.054**	**0.099**		
	[0.549]	**[0.060]**	**[0.111]**	-0.371	17,154
	(0.006)	(0.019)	(0.035)		
Oral	0.537	0.024	0.044		
	[0.546]	[0.047]	[0.086]	-0.311	3,275
	(0.011)	(0.065)	(0.119)		

Sample of children with dyslexia score within a 7 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 parentheses below. Mean of the outcome is reported for children with a dyslexia score above the cutoff. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).



(a) Academic well-being

(b) Social well-being

(c) Support and inspiration

(d) Calm and order

Figure E1. : Well-Being Subscales

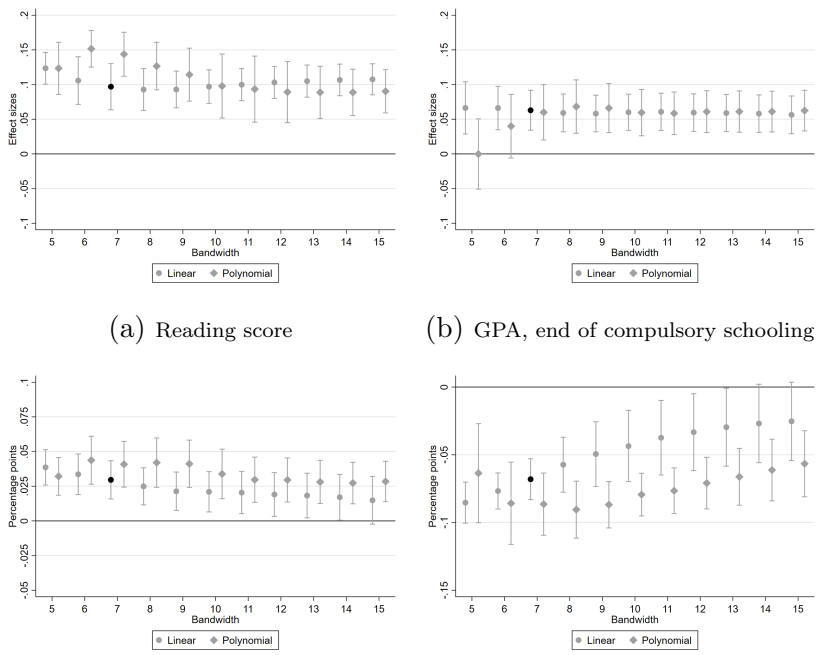
The figure presents bias-corrected reduced-form estimates along with robust confidence intervals. The estimates are based on a sample of children with dyslexia scores within a 7-point average range around the cutoff. The figure shows the four subscales underlying the school well-being composite, derived from the annual well-being survey conducted from Grade 4 to 9 in public schools.

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	**0.094**	**0.085**	**0.112**	**0.129**	**0.093**	**0.085**
	[0.097]	**[0.083]**	**[0.116]**	**[0.146]**	**[0.098]**	**[0.085]**
	(0.017)	(0.016)	(0.015)	(0.012)	(0.019)	(0.016)
Effective # of Obs.	15,142	15,142	11,192	13,279	13,844	16,780
GPA	**0.055**	**0.059**	**0.060**	**0.067**	**0.049**	**0.039**
	[0.063]	**[0.068]**	**[0.066]**	**[0.067]**	**[0.061]**	**[0.051]**
	(0.015)	(0.029)	(0.017)	(0.019)	(0.021)	(0.016)
Effective # of Obs.	16,638	16,638	14,614	19,989	15,170	18,444
Completed HS before age 20	**0.023**	**0.027**	**0.027**	**0.039**	**0.025**	0.008
	[0.030]	**[0.034]**	**[0.032]**	**[0.041]**	**[0.034]**	**[0.015]**
	(0.007)	(0.006)	(0.007)	(0.008)	(0.014)	(0.010)
Effective # of Obs.	12,354	12,354	10,808	13,745	11,296	13,745
NEET at age 20	**-0.053**	**-0.052**	**-0.070**	**-0.077**	**-0.038**	**-0.023**
	[-0.067]	**[-0.066]**	**[-0.083]**	**[-0.081]**	**[-0.056]**	**[-0.038]**
	(0.008)	(0.009)	(0.009)	(0.015)	(0.017)	(0.017)
Effective # of Obs.	6,257	6,257	4,582	5,453	5,724	7,007

Sample of children with dyslexia score within a 7 point average around the cutoff. Each cell reports the reduced-form coefficient from a separate local-linear regression with a triangular kernel of the outcome listed in the row, unless otherwise noted. Column 1 is the main specification and includes the covariates shown in Table D2; column 2 omits them. Column 3 uses the optimal bandwidth of Calonico, Cattaneo and Titiunik (2014), which differs by outcome. Column 4 uses a second-degree polynomial in the test score. Column 5 excludes observations with a test score exactly at the cutoff. Column 6 uses a rectangular kernel. Bias-corrected estimates are listed in square brackets and robust standard errors in parentheses below. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).



(a) Reading score

(b) GPA, end of compulsory schooling

(c) Completed HS before age 20

(d) NEET at age 20

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 result.

Suggestive Mechanisms

Table G1—: Heterogeneous Treatment Effect by Grade at Diagnosis

	Estimate (1)	BC estimate (2)	Robust S.E. (3)	Mean (4)	N (5)
A. Reading score					
- Elementary school (Grades 3–6)	**0.097**	[**0.102**]	(0.019)	-0.661	13,179
- Lower secondary education (Grades 7–8)	0.043	[0.025]	(0.034)	-0.632	1,963
B. GPA					
- Elementary school (Grades 3–6)	**0.050**	[**0.061**]	(0.031)	-0.536	13,028
- Lower secondary education (Grades 7–8)	0.073	[0.064]	(0.048)	-0.576	3,610
C. Completed high school before age 20					
- Elementary school (Grades 3–6)	**0.032**	[**0.041**]	(0.006)	0.186	8,612
- Lower secondary education (Grades 7–8)	0.002	[0.003]	(0.017)	0.191	3,742
D. Not in employment or education at age 20					
- Elementary school (Grades 3–6)	**-0.073**	[**-0.092**]	(0.018)	0.141	3,075
- Lower secondary education (Grades 7–8)	**-0.034**	[**-0.044**]	(0.014)	0.135	3,182

Sample of children with dyslexia score within a 7 point average around the cutoff, divided by grade at the time of diagnosis (elementary school: Grades 3–6; lower secondary education: Grades 7–8). Columns 1-5 report reduced-form results from separate local-linear regressions similar to the main equation with the 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 variable 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 (Calonico et al., 2014; 2019).

Table G2—: Heterogeneous Treatment Effect by Gender and SES

		Sex		SES	
	All	Boys	Girls	Low	High
	(1)	(2)	(3)	(4)	(5)
Reading score	**0.094**	**0.085**	**0.105**	0.076	**0.105**
	[0.097]	**[0.079]**	**[0.121]**	[0.079]	**[0.107]**
	(0.017)	(0.019)	(0.045)	(0.051)	(0.028)
Effective # of Obs.	15,142	8,424	6,718	7,688	7,437
GPA	**0.055**	**0.062**	**0.047**	**0.054**	**0.054**
	[0.063]	**[0.070]**	**[0.055]**	**[0.063]**	**[0.061]**
	(0.015)	(0.013)	(0.020)	(0.024)	(0.022)
Effective # of Obs.	16,638	9,176	7,462	7,805	8,815
Completed high school before age 20	**0.023**	**0.021**	*0.023*	**0.016**	**0.032**
	[0.030]	**[0.029]**	*[0.028]*	**[0.023]**	**[0.038]**
	(0.007)	(0.007)	(0.015)	(0.012)	(0.010)
Effective # of Obs.	12,354	6,879	5,475	6,101	6,239
Not in employment or education at age 20	**-0.053**	**-0.036**	**-0.074**	**-0.031**	**-0.066**
	[-0.067]	**[-0.044]**	**[-0.097]**	**[-0.037]**	**[-0.089]**
	(0.008)	(0.009)	(0.016)	(0.011)	(0.013)
Effective # of Obs.	6,257	3,605	2,652	3,170	3,081

Sample of children with dyslexia score within a 7 point average around the cutoff. The sample is stratified by sex and by socioeconomic status (SES), with low and high SES defined as below and above the sample median of a combined index of parental education and income rank. Each cell reports the reduced-form coefficient from a separate local-linear regression with a triangular kernel of the outcome listed in the row. Bias-corrected estimates are listed in square brackets and robust standard errors in parentheses below. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).

Table G3—: Year-grades

	First-stage (1)	Reduced form (ITT) (2)	LATE (3)	Mean outcome (4)	Effective obs. (5)
Year-marks	**0.540**	*0.019*	*0.035*	-0.490	17,472
	[0.546]	*[0.021]*	*[0.038]*		
	(0.006)	(0.012)	(0.022)		
Danish	**0.540**	**0.057**	**0.104**	-0.606	17,413
	[0.546]	**[0.060]**	**[0.111]**		
	(0.006)	(0.014)	(0.025)		
Math	**0.540**	**0.048**	**0.088**	-0.393	17,416
	[0.546]	**[0.049]**	**[0.091]**		
	(0.006)	(0.015)	(0.027)		
English	**0.539**	0.013	0.024	-0.598	17,254
	[0.546]	[0.014]	[0.026]		
	(0.006)	(0.019)	(0.035)		
Science	**0.541**	**0.018**	**0.033**	-0.325	16,920
	[0.547]	**[0.026]**	**[0.048]**		
	(0.006)	(0.009)	(0.016)		

Sample of children with dyslexia score within a 7 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 parentheses below. Mean of the outcome is reported for children with a dyslexia score above the cutoff. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (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 0 (1/0)	*-0.001*	*[-0.001]*	(0.001)	0.001	19,131
- Year 1 (1/0)	-0.003	[-0.003]	(0.002)	0.005	19,080
- Year 2 (1/0)	-0.005	[-0.006]	(0.003)	0.010	18,528
- Year 3 (1/0)	-0.005	[-0.003]	(0.002)	0.020	16,686
B. Private schooling					
- Year 0 (1/0)	**-0.004**	**[-0.005]**	(0.002)	0.128	19,131
- Year 1 (1/0)	**0.006**	**[0.006]**	(0.003)	0.132	19,080
- Year 2 (1/0)	0.002	[0.003]	(0.004)	0.138	18,528
- Year 3 (1/0)	**-0.011**	**[-0.010]**	(0.004)	0.145	16,686

Sample of children with dyslexia score within a 7 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 parentheses below. Mean of the outcome is reported for children with a dyslexia score above the cutoff. **Bold** (*italic*) indicates significance at 5% (10%) based on robust confidence intervals centered on bias-corrected estimates (Calonico et al., 2014; 2019).