



LUND UNIVERSITY

Adolescent Mental Health: Impact of Introducing Earlier Compulsory School Grades

Linder, Anna; Gerdtham, Ulf-Göran; Heckley, Gawain

2023

Document Version:
Other version

[Link to publication](#)

Citation for published version (APA):

Linder, A., Gerdtham, U.-G., & Heckley, G. (2023). *Adolescent Mental Health: Impact of Introducing Earlier Compulsory School Grades*. (Working Papers; No. 2023:2).

Total number of authors:
3

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2023:2

Department of Economics
School of Economics and Management

Adolescent Mental Health: Impact of Introducing Earlier Compulsory School Grades

Anna Linder
Ulf-G Gerdtham
Gawain Heckley

March 2023
Revised: November 2023



LUND
UNIVERSITY

Adolescent mental health: Impact of introducing earlier compulsory school grades

Anna Linder^{a,b,*}, Ulf-G Gerdtham^{a,b,c}, Gawain Heckley^{a,b}

a Health Economics Unit, Department of Clinical Sciences, Lund University, Malmö, Sweden

b Centre for Economic Demography, Lund University, Lund, Sweden

c Department of Economics, Lund University, Lund, Sweden

*Corresponding author: Anna Linder, Health Economics Unit, Lund University, Medicon Village 301:5, Nyttänkargatan 4, 223 63 Lund, Sweden. Tel: +46 72 467 48 80. E-mail: anna.linder@med.lu.se

The prevalence of mental illness is increasing among adolescents in many developed countries, and factors such as increased academic pressure and earlier introduction of performance evaluation may in part explain this trend. We study the effect of introducing earlier grades in compulsory school on adolescent mental health in Sweden. To do so, we exploit a grading reform in Swedish schools in which grades were introduced in 6th grade instead of 8th grade. Controlling for potential age effects in a difference-in-discontinuities setup, we show that girls who are exposed to earlier grades are more likely to be diagnosed with depression or anxiety by the end of compulsory school. We do not find similar effects among boys. These results indicate that the mental well-being of girls may be particularly influenced by educational assessments such as grades during early years, which suggests that potential health effects of different grading systems should be considered.

JEL codes: **I10, I21, I28,**

Key words: education policy, school grades, mental health, human capital development

Financial support for the study has been received from the Public Health Agency of Sweden (03478-2019). Funding is also gratefully acknowledged from the Swedish Research Council (DNR 2019-06292), Crafoord Foundation (dnr 20190685, dnr 20211004) and Tom Hedelius and Jan Wallander foundation (P2017-0081:1) (Heckley). Gerdtham is grateful to Region Skåne for financing part of his professorship. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. We thank Matthew Collins and Jonas Lundstedt for their helpful input to the paper. The data used in this paper comes from the Swedish Interdisciplinary Panel (SIP), administered by the Centre for Economic Demography, Lund University, Sweden. The study is approved by the Regional Ethics Committee in Lund (2012/627) and the Swedish Ethical Review Authority (2021-02630).

1. Introduction

A growing prevalence of adolescent mental illness is well documented across several high-income countries, including Sweden (Potrebny, Wiium et al. 2017). At the same time, there has been a trend of increasing school accountability in Sweden and other high-income countries, which involves ranking schools and students based on student performance measures (Figlio and Loeb 2011, Lundahl, Hultén et al. 2017). This trend is accompanied by a general shift towards increased top-down educational accountability within nations, resulting in, for example, increased high-stakes testing, assessment, and grading of children in schools (Figlio and Loeb 2011, Lingard, Martino et al. 2013). Most economies are also becoming more and more skill-intense, increasing the importance of school performance for future labor market opportunities, and potentially also increasing self-perceived school stress and mental health problems among adolescents (Högberg 2021). These trends suggest that increased attention to the mental health impacts of different assessment systems in schools is warranted.

Understanding the role of school assessment in the development of mental illness is important beyond the immediate health and welfare concerns for those affected. Mental health is an important input in the human capital accumulation process (Currie and Stabile 2006, Currie and Stabile 2007, Ding, Lehrer et al. 2009, Currie, Stabile et al. 2010, Fletcher 2010) and there are several mechanisms by which school grades can impact both mental health and cognitive skills. Grades provide information on current achievement and progress. They can also provide a tool for enhanced learning by potentially motivating students to achieve success or avoid failure (Klapp, Cliffordson et al. 2016). Through these functions, grades have the potential to impact higher education opportunities and the development of cognitive human capital. These impacts may then lead to indirect effects on well-being or mental health (Heckman 2012). However, there may also be a more direct effect on human capital in terms of health. Theories in symbolic interactionism claim that children's perceptions about themselves are based on peers', parents', or teachers' appraisals about them and that positive or negative feedback shapes their self-image and promotes or inhibits mental health, which may introduce a direct link between grades (feedback) and mental health (Gustafsson, Allodi et al. 2010). These direct impacts on mental health, may then affect future cognitive skill development.

In this paper we study the effect of earlier grading in schools in Sweden on the mental health of adolescents. We exploit a reform to the grading system in Sweden in which grades were introduced at an earlier age, in grade 6 instead of grade 8. Compared to other Nordic countries, the Swedish grading system introduces grading of students' performance at a relatively early age (in grade 6, around the age of 13) (Lundahl, Hultén et al. 2017). Moreover, since grades in compulsory schools in Sweden are used for purposes such as admission and selection to upper secondary schools, they are perceived as high-stake which is not comparable with most other countries using early grading (Lundahl, Hultén et al. 2017). The reform was introduced in the fall of 2012, which means that children born up until the year 1998 received grades according to the old system, in 8th grade, while children born in 2000 and thereafter were exposed to grades for the first time according to the new system, in 6th grade. Children born in 1999 entered 7th grade in the year of the reform and were the only cohort that received grades for the first time in that school year. The reform provides a situation where the age when children receive their first grade is arbitrary depending on whether the child is born before or after a year-end of 2000 (school entry cutoff). We can use this discontinuity in exposure to grades at different ages to estimate the causal effect of earlier grading on mental illness.

There are, however, circumstances that prevent us from using standard regression discontinuity design. Previous literature has established that school-starting age affects academic performance and subsequent human capital accumulation, and highlights mechanisms that are of relevance to our empirical setup (Fredriksson and Öckert 2014). At school start, and at first exposure to grades in schools, children who are born in January are almost one year older than children born in December. Since mental illness increases with age in adolescence and young adulthood, older children will have a higher probability of mental illness since the follow-up period is the same for all children in a cohort. Moreover, being relatively older or younger than peers may also impact mental health through different responses to earlier grades. If we expect younger children to be more sensitive to grades it would mean that children born later in a year will have a higher probability of mental illness. To account for such absolute and relative age effects, we exploit two different sources of variation in a so-called, Difference in Discontinuities (Diff-in-Disc) setup, see Grembi, Nannicini et al. (2016) and Collins and Lundstedt (2021) for examples of studies with the same method. The discontinuity in mental illness at the school entry cutoff for cohorts that have been exposed to the reform (earlier grades) is differenced out by the discontinuity in

mental illness for control cohorts that is related only to absolute and relative age effects. Left is the causal effect of earlier grades on mental illness.

We utilize Swedish individual-level data of all children born between the years 1992 and 2000. The results show that earlier grades are related to increased mental illness by the end of compulsory school. Girls who are exposed to one year earlier grades are 0.63, 0.44, and 0.36 percentage points more likely to receive a diagnosis for any internalizing disorder, depression, or anxiety, respectively, which relates to increases in the probability of diagnosis by roughly 50-70 percent depending on the outcome. We do not find similar effects among boys.

Our findings are in line with previous results by Högberg, Lindgren et al. (2021), who find that earlier grades have a negative impact on self-reported school-related stress, academic self-esteem, psychosomatic symptoms, and life satisfaction, especially among girls. However, this research relies on self-reported measures of well-being and therefore changes may just reflect a difference in willingness to report. Furthermore, their results also vary by choice of control group, suggesting a degree of caution is warranted. To our knowledge this is the only other available research on extensive margin school grading effects (grades or no grades) on mental health. The reason the literature on extensive margin grading effects is limited may be that grading systems usually are determined at the national level and reforms are relatively uncommon (Lundahl, Hultén et al. 2016), which means that there is limited variation in whether and when school grading is introduced between comparable populations. Sweden is an exception, where repeated reforms to the grading system have given rise also to a few studies on the effect of grades on measures of cognitive achievement and educational achievement. Using a reform to the grading system in Sweden in the 1980s, in which written reports of grades at the end of lower and middle school were abandoned, Sjögren (2010) found evidence of a subsequent reduced number of years of schooling for girls, and Klapp (2015) found that this reform reduced subsequent compulsory school grades and reduced the probability of graduating from upper secondary school. Another study which is on the impact of arbitrarily inflated grades on mental health shows that over-grading, that is, getting a grade that is higher than actual skills or achievements, leads to a reduced probability of depression- and anxiety-related diagnoses and prescriptions among young women in Sweden (Linder et al. 2023). Most of these studies finds that the impacts of grades differ by sex, as well as other child characteristics, for example, own skills and parents' education (Sjögren 2010, Klapp

2015, Linder, Nordin et al. 2023), which is the general finding also in related literature investigating test-based assessment of students (Dee and Jacob 2011, Whitney and Candelaria 2017). These findings implies that girls and young women are particularly receptive to the consequences of grades, with potentially unintended effects on their health and human capital accumulation, which may be related to the gap in mental health problems, especially regarding depression and anxiety, that appears between boys and girls in school-age and persists through life.

Our paper contributes to a growing literature on the consequences of performance feedback in the context of grade-based assessment and to understanding the drivers behind increasing levels of mental illness. We provide novel evidence of how changes in a school grading policy can impact medically diagnosed mental illness. Changes in the school system that potentially have no impact on the skills accumulation of students can still potentially have meaningful impacts on health capital, a concern by itself, but of further concern since such health impacts may affect future development (Currie and Stabile 2006, Currie and Stabile 2007, Ding, Lehrer et al. 2009, Currie, Stabile et al. 2010, Fletcher 2010). We use a credible identification strategy to isolate the impact of a grading reform to understand how earlier grading affects the mental health of adolescents. The use of administrative data on mental health-related diagnoses means that we identify an objective change in mental illness, as opposed to self-reported measures of mental health. Large sample sizes also allow detailed heterogeneity analysis to understand the mechanisms at play.

In the next section, we briefly describe the background to the Swedish schooling system with regards to the grading system and the reform that we exploit to identify exogenous variation in age at first grading in school. In sections three and four, we present our data and empirical strategy. Section five presents the results, including sensitivity tests and investigation of potential timing of the effects. In section six we provide several robustness checks such as manipulation-, covariate balance-, and placebo-tests, and in section seven, we investigate heterogeneity in the effects based on parents socioeconomic and migration background. Finally, in sections number eight and nine, we discuss and conclude our findings.

2. The 2012 Grading Reform

At the beginning of the 1990s, several reforms to the Swedish schooling system were implemented that were still in place during the study period we consider: a decentralization that shifted the responsibility of primary and secondary schools from the central state to the local municipalities; a deregulation of school choice and financing meant that students are now able to choose schools outside of their catchment area including private schools that after the reform could receive public funding through a voucher system; and finally, the grading system was reformed from a norm-referenced system to a goal-oriented system. The reforms were followed by a period of substantial grade inflation in Swedish upper secondary schools that was not reflected in increased student performance (Wikström and Wikström 2005, Vlachos 2011). Swedish students' performance worsened during the first decade of the 2000s according to international assessments such as the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) (The Public Health Agency of Sweden 2018). At the same time, the share of students eligible to progress from compulsory school to upper secondary school decreased (The Swedish Government 2009). In consequence, the Swedish Government presented additional reforms to the schooling system. A new curriculum was introduced in compulsory schools in Sweden in the fall of 2011 (Lgr11) (The Swedish National Agency for Education 2011). The main purpose of the reform of the curriculum was to revise the syllabus and clarify the learning outcomes in the different courses to make them easier to understand. The new curriculum affected all children born as of January 1st 1996. In connection with the Lgr11, a more informative grading scale was introduced, transforming a three-level grading scale (pass; pass with distinction; and pass with special distinction) to a six-level grading scale, A–F, where A–E are different pass grades and F reflects the failure to reach the learning outcomes set for level E. The reform of the grading scale affected children entering 8th grade in the fall of 2011, that is, children born as of January 1st in 1997 (children entering 9th grade in 2011 received grades according to the old system).

Following these, changes regarding earlier grading in compulsory school, from grade 6 instead of grade 8, were introduced in the fall of 2012. Children entering 6th and 7th grade at this time received grades for the first time in that school year, which means that children born in 1999 received grades from 7th grade, and children born as of January 1st in 2000 received grades from 6th grade. Term grades are given at the end of each semester, and a final grade is

given at the end of the spring semester in grade 9. The grade shall reflect the student's knowledge based on the specific grade criteria at the time the grade is set. This means that teachers can choose to include achievements from the previous semester if they relate to the current assessment criteria, but the student's performance from the current semester has the most weight. It is the final grade at the end of compulsory school that is used for selection into upper secondary school. Since there is a limited connection between term grades and the final grade, especially for the first years of grading, the term grades are generally not considered "high-stakes". Nevertheless, reports have shown that for many children also the grades before the final grade are perceived as high-stake (The Swedish National Agency for Education 2017).

The main purpose of introducing grades at an earlier age was to improve the achievement of learning outcomes by earlier detection of students needing extra support (The Swedish National Agency for Education 2017). Teachers report that the reform indeed facilitated the identification of students in need of extra support, but nothing indicates that the reform actually increased the support for these students (The Swedish National Agency for Education 2017). An additional purpose of the grading reform was to transfer information to the parents on their child's skills and performance. If parents were more informed about their children's education they can also participate in and support their development more actively and efficiently. According to a study on parents' experiences of the grade reform, the parents experienced large variations in both amount and quality of the information on their children's performance. This variation in information was linked both to differences in the provided information by the school/teachers and linked to parents' interest, effort, and ability, which was linked to socioeconomic background (Brismark, Hultin et al. 2016). Throughout the rest of the paper, we refer to this reform as the earlier grades reform.

3. Data

We use data from the Swedish Interdisciplinary Panel (SIP) administered by the Centre for Economic Demography at Lund University, Sweden. SIP contains the entire Swedish population born between 1973 and 2016, and consists of data from various administrative databases from, amongst others, Statistics Sweden and the Swedish National Board of Health and Welfare. The study population comprises children born between the years 1992 and 2000. The nature of the Diff-in-Disc strategy requires that we have children distributed in cohorts that are re-centered around a year-end. We measure the impact of the earlier grades

reform on mental health in two reform cohorts. The first cohort consists of children born six months before and after January 1st in 1999, where those born before the turn of the year receive grades from school year 8 and those born after the year-end receive grades from year 7. The second cohort consists of children born six months before and after January 1st in 2000 (where those born before the turn of the year receive grades from school year 7 and those born after the year-end receive grades from year 6) (see figure 1). Corresponding cohorts born around January 1st in the years 1993, 1994, and 1998 are included in the study sample as controls. The cohorts born around January 1st in 1995, 1996, and 1997 are not included as controls since these cohorts were exposed to other grading reforms¹.

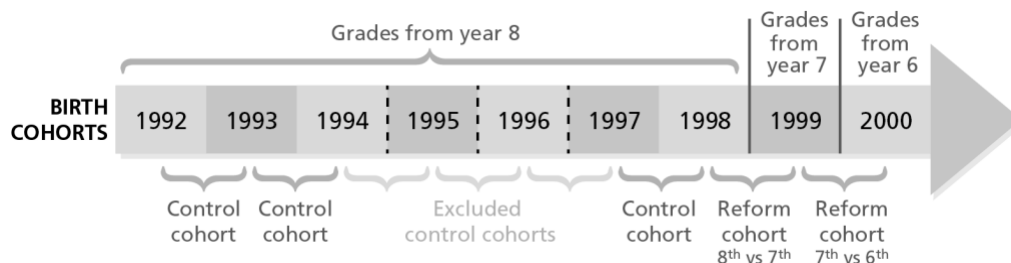


Figure 1 The Grading System by Birth Cohort

Note: This figure illustrates grading in the analysis sample birth cohorts. The birth cohorts represent all children born from January 1st to December 31st in a calendar year and the re-centered analysis cohorts (below arrow) represent children born from July 1st in a year to June 30th in the next year.

The relevant birth cohorts are drawn from the Register of the Total population and linked to their parents (biological and adoptive) through the Multigenerational register. Since we measure the effect of earlier grades (the introduction of grades in 6th, 7th, or 8th grade) on mental health outcomes in the year when the child starts the final year of compulsory school (9th grade), we exclude those who immigrated to Sweden later than grade 5 and those who emigrated from Sweden before end of compulsory school. The final sample comprises 524,093 individuals.

To this sample, we match background information on demographic, migration, and socioeconomic factors, as well as inpatient and outpatient diagnoses from the National Patient registers. Our mental health outcome measures are dummies indicating the probability

¹ Lgr11 (compulsory school) and Gy2011 (upper secondary school) were introduced in 2011.

of mental illness, that is at least one diagnosis² in inpatient or specialized outpatient care for internalizing disorders and substance use disorders, the most common psychiatric diagnoses with onset in adolescence and young adulthood in Sweden (The National Board on Health and Welfare, 2020). In general, internalizing disorders are emotional or behavioural mental problems that are internalized – kept to themselves – by the individual, often characterized by a change in mood or affect, for example, depression, a neurotic reaction to an environmental or stressful situation, or a general manifestation of anxiety. In our outcome measure *internalizing disorder* we include mood (affective) disorders (ICD-10 codes F30-39) and neurotic, stress-related, and somatoform disorders (ICD-10 codes F40-48). Separate outcomes for *depression* (depressive episode and recurrent depressive disorder, ICD-10 codes F32-33), *anxiety* (phobic anxiety disorders, other anxiety disorders, and obsessive-compulsive disorders, ICD-10 F40-42), and *stress* (reactions to severe stress and adjustment disorders, ICD-10 F43), are also investigated. Substance use disorder diagnoses are attributable to acute intoxication, harmful use, dependence syndromes, and other conditions related to the use of one or more psychoactive substances such as alcohol, narcotics, or tobacco. Our outcome *substance use disorders* include mental and behavioural disorders related to psychoactive substance use (ICD-10 codes F10-19), and we also investigate separate outcomes for *alcohol-related disorders* (ICD-10 F10) and *narcotics-related disorders* (ICD-10 codes F11-16, F18, F19).

Migration background is defined by two variables: 1) *foreign-born*; and 2) *foreign-born parents* (second-generation immigrant). Parents' education level is defined by three variables: 1) *low-educated parents* if both parents have only compulsory schooling; 2) *moderately educated parent* if either of the parents has more than compulsory schooling but none of them have completed tertiary schooling, and 3) *high educated parent* if at least one parent has completed tertiary education. Data on *parents' income* is household income (equivalized by household composition) extracted from the parents in the year when the child turns 10 years, that is, before exposure to the reform.

² Primary diagnosis, as well as the top three secondary diagnoses, are included.

Table 1: Descriptive Statistics

	(1)	(2)	(3)	(4)
	Girls		Boys	
	Reform cohorts	Control cohorts	Reform cohorts	Control cohorts
Any internalizing diagnosis	0.0166 (0.128)	0.0153 (0.123)	0.0065 (0.0804)	0.0063 (0.0794)
- Depression	0.0082 (0.0900)	0.0073 (0.0851)	0.0028 (0.0525)	0.0026 (0.0505)
- Anxiety	0.0094 (0.0966)	0.0071 (0.0843)	0.0036 (0.0597)	0.0031 (0.0553)
- Stress	0.0021 (0.0459)	0.0025 (0.0495)	0.0008 (0.0285)	0.0008 (0.0284)
Any substance use diagnosis	0.0019 (0.0432)	0.0043 (0.0654)	0.0017 (0.0409)	0.0036 (0.0601)
- Alcohol	0.0012 (0.0343)	0.0036 (0.0602)	0.0008 (0.0289)	0.0028 (0.0531)
- Narcotics	0.0008 (0.0289)	0.0008 (0.0282)	0.0009 (0.0301)	0.0009 (0.0305)
Parents income (SEK)	267,093 (288.2)	228,636 (232.2)	265,131 (246.7)	228,988 (271.4)
Low educated parents	0.0412 (0.199)	0.0408 (0.198)	0.0414 (0.199)	0.0405 (0.197)
Moderately educated parent	0.4859 (0.500)	0.5287 (0.499)	0.4892 (0.500)	0.5304 (0.499)
High educated parent	0.4730 (0.499)	0.4304 (0.495)	0.4695 (0.499)	0.4290 (0.495)
Foreign-born	0.0876 (0.283)	0.0712 (0.257)	0.0880 (0.283)	0.0682 (0.252)
Foreign-born parents	0.1091 (0.312)	0.0940 (0.292)	0.1066 (0.309)	0.0958 (0.294)
Observations	91,958	162,943	98,076	171,116

Notes: This table shows the mean probability of the prespecified mental illness outcomes and characteristics on parental SES (income and education level), as well as migration background, among girls and boys separately in the reform and control cohorts. The reform sample consists of all children born in Sweden six months before and after January 1st in 1999 and 2000 and the control sample consists of cohorts born six months before and after January 1st in 1993, 1994, and 1998 (see data section for specific information about the sample). Standard deviations in parenthesis. The mental health outcomes include the following ICD-10 codes: Internalizing disorders diagnosis F30-39 and F40-48); depression F32-33; anxiety F40-42; stress F43; substance use disorders F10-19; alcohol F10; and narcotics F11-16, F18 and F19. Parents income is the mean of the mother's and father's household income (equivalized by household composition) in the year the child turns 10 years.

4. Empirical Strategy

Our empirical strategy utilizes the Regression Discontinuity (RD) framework. In the following standard RD model, the sample is restricted to the two re-centered reform cohorts born around January 1st in the years 1999 and 2000:

$$y_i = \alpha_0 + \mathbf{x}'_i \boldsymbol{\alpha}_1 + \rho_{RD} D_i + (D_i * \mathbf{x}_i)' \boldsymbol{\alpha}_2 + \varepsilon_i \quad (1)$$
$$D_i(x_i) = \mathbb{I} [x_i \geq 0]$$

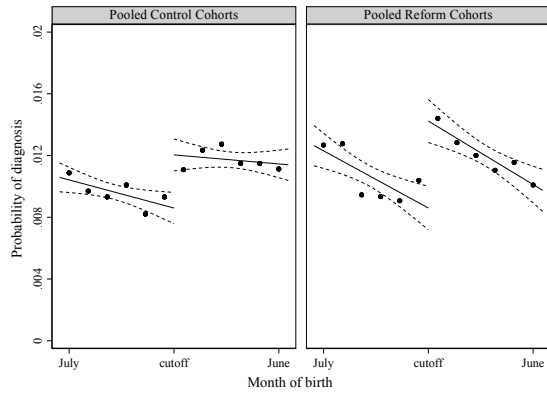
Where \mathbf{x}_i is a vector of the running variable of increasing order polynomials measured as month of birth, re-centered around the school entry cutoff, the 1st of January. y_i is the outcome of interest, mental illness for individual i . ρ_{RD} is the coefficient for the treatment indicator D_i , which is a deterministic function of the first order polynomial of x_i and equal to 1 if exposed to the reform (born in January to June). ρ_{RD} is the parameter of interest and captures the effect of being exposed to the reform. Under the following assumptions, the RD estimator ρ_{RD} can be interpreted as the treatment effect of one year earlier grading on the probability of mental illness:

Assumption I: All potential outcomes are continuous in x at the cutoff

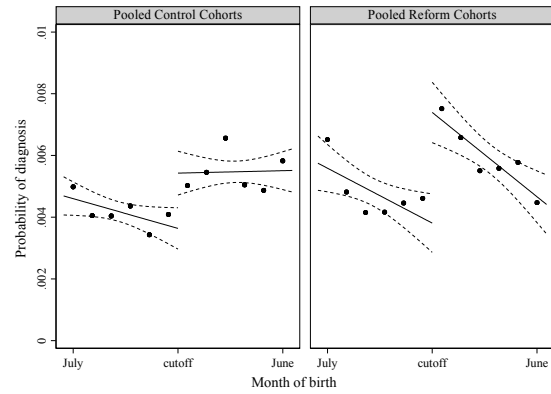
Assumption II: Full compliance, being born in January to June means exposure to the reform

In figure 2, the mean probabilities of the mental health outcomes are plotted by month of birth. The plots are provided separately for the pooled reform cohorts (those born around January 1st in 1999 and 2000) and the pooled control cohorts (those born around January 1st in years 1993, 1994 and 1998). The discontinuity graphs reveal large shifts at the reform cohort cut-offs for some of the outcomes, especially for internalizing disorders, but also for depression and anxiety separately, implying that earlier grading increases the probability of these disorders. However, there are shifts at the cut-off also for the pooled control cohorts. Children born early in the year are up to one year older than children born late in the year when they are exposed to grades for the first time. Relative age effects, being older or younger than your peers, may impact the response to earlier grades. Moreover, absolute age effects, that the probability of mental illness increases with age in adolescence, means that the probability of mental illness may be higher for the older children. The shifts at the school cut-

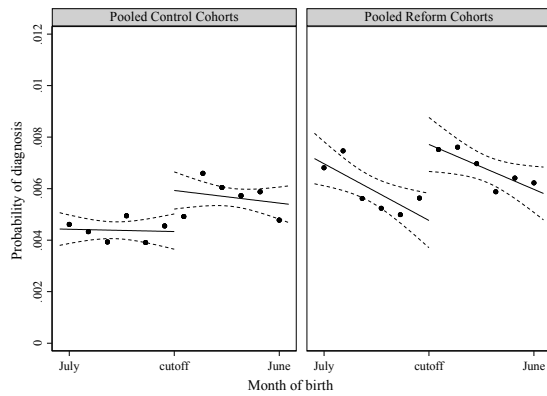
off for the control cohorts implies that there are age effects on mental illness, which may be of relevance also for other potential outcomes. Since this is a violation of the first identifying assumption, we need to adjust our baseline model.



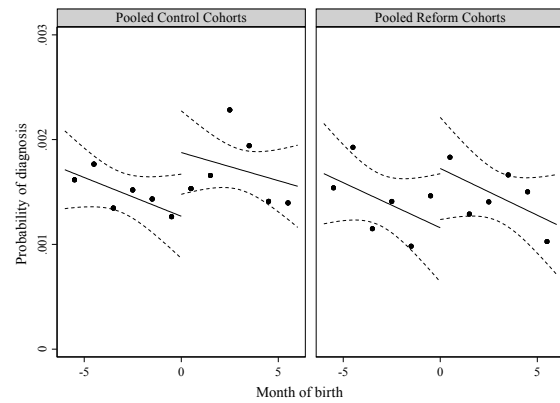
a: Internalizing Disorder



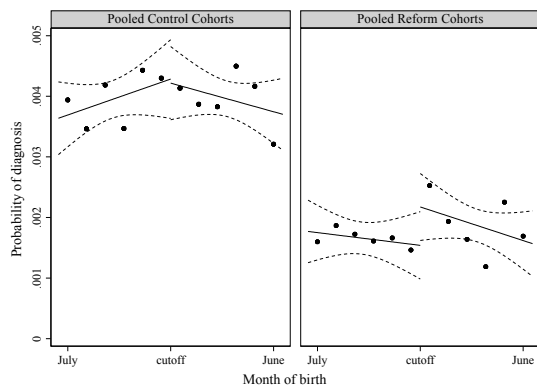
b: Depression



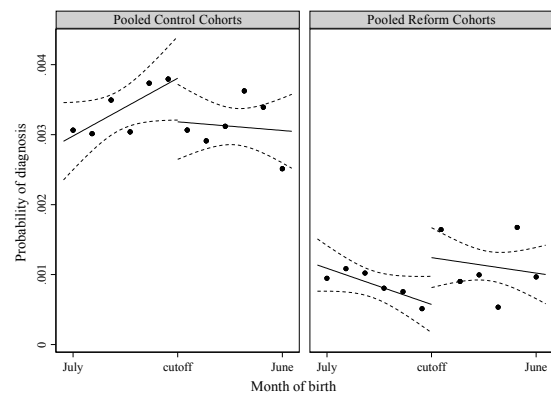
c: Anxiety



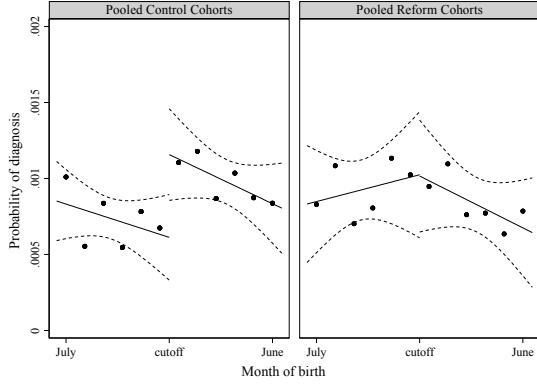
d: Stress



e: Substance Use Disorder



f: Alcohol



g: Narcotics

Figure 2 Discontinuity Graphs – Mental Illness Outcomes

Note: In this figure, the mean probability of the mental illness outcomes and fitted linear predictions of diagnosis (including 95 % clustered along birth month and year confidence intervals) are plotted by month of birth (July to June). The plots are provided separately for the pooled control cohorts (those born around January 1st in years 1993, 1994 and 1998), and the pooled reform cohorts (those born around January 1st in 1999 and 2000). The shift at the control cohort cut-offs reflects age (month of birth) effects while the shift at the reform cohorts cut-offs reflects age effects *and* the potential effect of one year earlier grading in schools on mental illness.

To correct for the discontinuity in x at the cutoff, we include the series of control cohorts in the regression model. Conceptually, this means combining the basic RD model with a Difference-in-Difference component in a so-called, Difference-in-Discontinuities (Diff-in-Disc) model (Grembi et al. 2016, Collins and Lundstedt 2021). By including the control cohorts, the discontinuity in potential outcomes at the school entry cutoff for cohorts that have been exposed to earlier grades is differenced out by the discontinuity for control cohorts that is related only to absolute and relative age effect. In the following Diff-in-Disc model,

$$y_i = \beta_0 + \beta_1 D_i + \mathbf{x}'_i \boldsymbol{\beta}_2 + (D_i * \mathbf{x}_i)' \boldsymbol{\beta}_3 + C_i \cdot \{\rho_{DD} D_i + \mathbf{x}'_i \boldsymbol{\beta}_4 + (D_i * \mathbf{x}_i)' \boldsymbol{\beta}_5\} + \pi_{C_i} + \eta_i, \quad (2)$$

$$C_i = \mathbb{I} [\text{reform cohort} = 1],$$

Where Y_i , D_i , and \mathbf{x}_i stem from Eq. 1. C_i indicates the value 1 if born in one of the two re-centered reform cohorts (around January 1st in 1999 and 2000), and π_{C_i} are birth cohort fixed effects. In the sensitivity analysis we also test if there is effect heterogeneity between the two reform cohorts, that is, if the response to earlier grades differs with age/school year. In this case the indicator function C_i equals 1 if the reform cohort is born either around January 1st in 1999 *or* around January 1st in 2000, excluding the other reform cohort from the analysis.

corresponding birth- and school cohorts (see figure 4). Less than 5 percent of children born in February – October are assigned to a different school cohort while 6 and 10 percent of children born in November and December are assigned to a late school cohort and 7 percent of children born in January are assigned to an early school cohort. Therefore, we cannot assume that all children born after January 1st in the reform cohorts actually are exposed to the reform and there may be selection in which children are exposed to the reform.

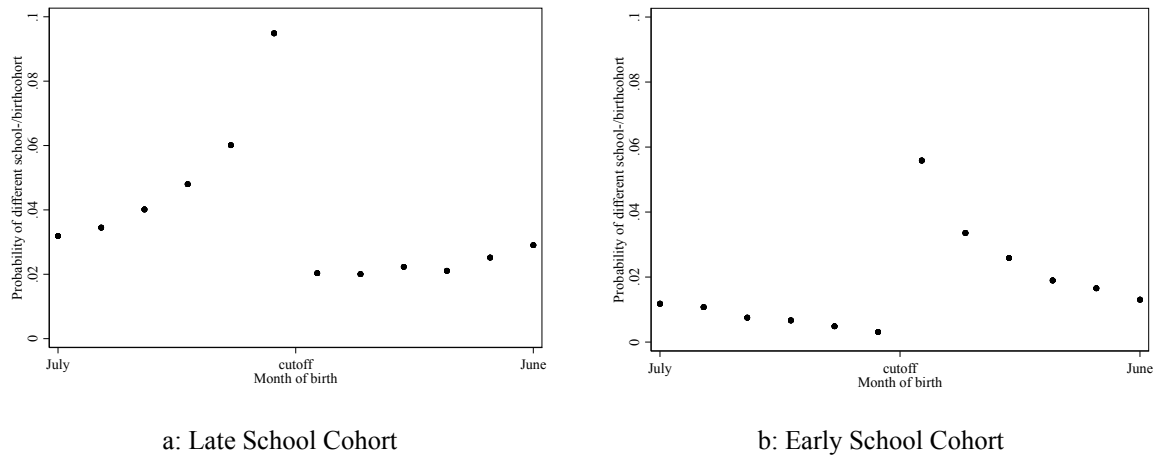


Figure 4: School-Cohort Assignment

Note: This figure shows mean probability of assignment to a late or early school-cohort not corresponding with one's birth-cohort (late school cohort means being with younger peers and early school cohort means being with older peers, in school), plotted in bins by month of birth running from July to June. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample).

To overcome this issue, we instrument exposure to the earlier grades using the school entry cutoff based on birth month and year in a fuzzy Diff-in-Disc. Fuzzy discontinuities are used within the RD framework when the probability, or the intensity, of treatment jumps at the cutoff, as opposed to treatment switching on and off as for the sharp RD. By instrumenting exposure to earlier grades by birth month and year in this way, we can isolate the variation in assignment to school cohort that is exogenous and use this unproblematic part of the variation to estimate an unbiased effect of earlier grades on mental illness. This means that Eq 2, where the actual treatment is unobserved, can be seen as the reduced form equation in an instrumental variables context. Consequently, in the following system of equations, Eq 3 represents the first stage which captures the jump in actual exposure to the reform induced by birth month and year and produces the variable \hat{S}_i , the fitted value indicating actual exposure to the reform which appears in Eq. 4, the two-stage least squares (2SLS) second stage.

$$S_i = \gamma_0 D_i + \mathbf{x}'_i \boldsymbol{\gamma}_1 + (D_i * \mathbf{x}_i)' \boldsymbol{\gamma}_2 + C_i * \{ \rho_{FS} D_i + \mathbf{x}'_i \boldsymbol{\gamma}_3 + (D_i * \mathbf{x}_i)' \boldsymbol{\gamma}_4 \} + \pi_{C_i} + \mu_i \quad (3)$$

$$Y_i = \delta_0 \hat{S}_i + \mathbf{x}'_i \boldsymbol{\delta}_1 + (\hat{S}_i * \mathbf{x}_i)' \boldsymbol{\delta}_2 + C_i * \{ \rho_{2SLS} \hat{S}_i + \mathbf{x}'_i \boldsymbol{\delta}_3 + (\hat{S}_i * \mathbf{x}_i)' \boldsymbol{\delta}_4 \} + \pi_{C_i} + \epsilon_i \quad (4)$$

Under two additional identifying assumptions our parameter of interest, the 2SLS parameter ρ_{2SLS} , can be interpreted as the local average treatment effect (LATE) of earlier grading on mental illness among children born close to the school cohort cutoff in the two reform cohorts.

Assumption IV: there is a strong first stage between the instrument and the exposure to the reform

Assumption V: the exclusion restriction, that the instrument (birth month and year) affects the outcome (mental illness) only through the reform (earlier grades) given our identification strategy

First stage estimates and joint significance F-test of the instruments are provided in the appendix (see table A1). Standard errors are clustered on birth month and year in all analyses, but we also provide results in the appendix with robust standard errors (see table A2). We test the robustness of our preferred specification (first order polynomial in the running variable) using 2nd and 3rd order polynomials of the running variable to allow for a more flexible relationship between birth month and mental illness (see table A3), as well as testing the use of alternative bandwidth including children born three months before and after the cutoff (see table A4).

5. Results

5.1 Main Results

The results are presented in table 2. Our main finding is that being exposed to grades one year earlier (in 6th vs 7th grade or in 7th vs 8th grade) is significantly associated with an increased probability of mental illness in the year entering the final year of compulsory school (9th grade). The effect is driven by internalizing disorders, mainly depression and anxiety, among

girls. The 2SLS estimates show that earlier grades are significantly associated with 0.63, 0.44, and 0.36 percentage points increase in the probability of receiving a diagnosis for any internalizing disorder, depression, or anxiety among girls, respectively (see table 2). The absolute effect estimates are low but given the lower mean probability of diagnosis at this age, these effects relate to increases in the probability of diagnosis by roughly 50-70 percent depending on the outcome. Our results also show that earlier grades increase the probability of alcohol disorders by 0.15 percentage points (a 65 percent increase) and decrease the probability of narcotics-related disorder by 0.06 percentage points (a 90 percent decrease) in the full sample. The absolute probabilities for these outcomes are initially very low resulting in less stable results compared to the internalizing disorders. Yet, all estimates for substance use diagnoses, also for girls and boys separately, develop in the same direction, indicating that there may be earlier grades responses increasing the probability of alcohol-related disorder and decreasing the probability of narcotics-related disorders among children. Since the results indicate that the impact of earlier grades on mental health differs between girls and boys, all analyses from now on are performed for girls and boys separately.

5.2 Sensitivity to Excluded Control Cohorts

In the baseline model, we investigate the impact of earlier grades in the two reform cohorts (those born around January 1st in 1999 and 2000) while controlling for age effects by including a set of control cohorts (those born around January 1st in years 1993, 1994, and 1998). The cohorts born around January 1st in 1995, 1996 and 1997 are not included as controls since they were exposed to other grading reforms (see data section). In this section, we investigate the sensitivity of our results to the excluded control cohorts (born around January 1st in 1995 – 1997) by including each of these cohorts as controls, separately and together.

Our main results of increased mental illness following earlier grades are robust to the inclusion of the excluded control cohorts (see table 3). The results for internalizing disorders, and for depression separately, among girls are very stable and similar to the main results, and the effects on anxiety and alcohol-related disorders are also in the same direction as the main results, even though the effect is not statistically significant in all specifications. The results indicate that the effects of earlier grading on mental illness become more pronounced by the inclusion of the 1996-cohort (exposed to the Lgr11 reform) which supports the decision to exclude this cohort from our main model.

Table 2: Main Results – The Effect of Earlier Grading on Mental Illness at the End of Compulsory School, 2SLS

VARIABLES	(1) Any internalizing diagnosis	(2) Depression	(3) Anxiety	(4) Stress	(5) Any substance use diagnosis	(6) Alcohol	(7) Narcotics
Full sample							
Earlier grades reform	0.00244* (0.00131)	0.00202** (0.000902)	0.00151* (0.000904)	-6.05e-05 (0.000461)	0.000796* (0.000463)	0.00148*** (0.000429)	-0.000643** (0.000251)
Constant	0.00944*** (0.000537)	0.00414*** (0.000333)	0.00552*** (0.000379)	0.00127*** (0.000184)	0.00288*** (0.000198)	0.00227*** (0.000171)	0.000718*** (0.000121)
Observations: 524,093							
Girls							
Earlier grades reform	0.00629*** (0.00210)	0.00441*** (0.00149)	0.00359** (0.00147)	0.000139 (0.000645)	0.00107 (0.000894)	0.00172** (0.000705)	-0.000596 (0.000496)
Constant	0.0132*** (0.000969)	0.00627*** (0.000700)	0.00784*** (0.000607)	0.00180*** (0.000239)	0.00379*** (0.000317)	0.00333*** (0.000257)	0.000646*** (0.000208)
Observations: 254,901							
Boys							
Earlier grades reform	-0.00124 (0.00143)	-0.000272 (0.000858)	-0.000467 (0.000962)	-0.000248 (0.000572)	0.000542 (0.000956)	0.00127 (0.000807)	-0.000696 (0.000490)
Constant	0.00599*** (0.000710)	0.00220*** (0.000406)	0.00339*** (0.000410)	0.000801*** (0.000251)	0.00201*** (0.000366)	0.00128*** (0.000303)	0.000780*** (0.000209)
Observations: 269,192							

Note: This table presents the results for the impact of earlier grades (in 6th vs 7th and 7th vs 8th grade) on the probability of mental illness in the year the child enters 9th grade in compulsory school. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (ρ_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample). Cluster along birth month and year robust standard errors in the parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 3: Sensitivity to Inclusion of Additional Control Cohorts

VARIABLES	(1) Any internalizing diagnosis	(2) Depression	(3) Anxiety	(4) Stress	(5) Any substance use diagnosis	(6) Alcohol	(7) Narcotics
Girls							
Earlier grades reform, incl 1995 cohort	0.00496**	0.00407***	0.00272*	-0.000203	4.00e-05	0.000861	-0.000828*
Observations: 308,464	(0.00205)	(0.00146)	(0.00141)	(0.000667)	(0.000972)	(0.000779)	(0.000484)
Earlier grades reform, incl 1996 cohort	0.00825***	0.00568***	0.00407***	4.06e-05	0.00116	0.00182***	-0.000563
Observations: 304,610	(0.00231)	(0.00162)	(0.00146)	(0.000650)	(0.000814)	(0.000647)	(0.000467)
Earlier grades reform, incl 1997 cohort	0.00642***	0.00457***	0.00298**	0.000262	0.000352	0.00112	-0.000657
Observations: 302,639	(0.00198)	(0.00141)	(0.00147)	(0.000631)	(0.000920)	(0.000730)	(0.000464)
Earlier grades reform, incl 1995-97 cohort	0.00678***	0.00513***	0.00291**	-7.95e-05	-5.14e-05	0.000796	-0.000775*
Observations: 405,911	(0.00209)	(0.00147)	(0.00139)	(0.000632)	(0.000860)	(0.000676)	(0.000449)
Boys							
Earlier grades reform, incl 1995 cohort	-0.00171	-0.000509	-0.000513	-0.000164	0.000288	0.00108	-0.000687
Observations: 325,801	(0.00137)	(0.000808)	(0.000913)	(0.000540)	(0.000922)	(0.000774)	(0.000484)
Earlier grades reform, incl 1996 cohort	-0.000685	0.000157	1.21e-06	-0.000160	0.00113	0.00165**	-0.000544
Observations: 321,664	(0.00143)	(0.000852)	(0.000957)	(0.000530)	(0.000927)	(0.000745)	(0.000492)
Earlier grades reform, incl 1997 cohort	-0.00130	-9.72e-05	-0.000491	-0.000273	0.000277	0.00102	-0.000907*
Observations: 319,364	(0.00136)	(0.000800)	(0.000921)	(0.000546)	(0.000863)	(0.000723)	(0.000480)
Earlier grades reform, incl 1995-97 cohort	-0.00124	-3.03e-05	-0.000200	-0.000147	0.000593	0.00123*	-0.000727
Observations: 428,445	(0.00133)	(0.000780)	(0.000891)	(0.000509)	(0.000843)	(0.000689)	(0.000473)

Note: This table presents the sensitivity of including additional control cohorts on the main results in Table 2. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (p_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000, as well as the additional included cohort(s) specified in the table (see data section for specific information about the sample). Cluster along birth month and year robust standard errors in the parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5.3 Sensitivity to Model Specification

In our baseline model, we estimate the effect of earlier grading, treating the effect of one year earlier grades in 6th vs 7th grade and 7th vs 8th grade as the same. In the sensitivity analysis we test if there is effect heterogeneity between the two reform cohorts, that is if the response to earlier grades differs between the reform cohort born around January 1st in 1999 and the reform cohort born around January 1st in 2000. This analysis shows that it is mainly the effect of earlier grading in 7th grade instead of 8th grade that appears to affect mental health, both for depression and anxiety among girls and for alcohol-related disorders among girls and boys (see figure 5). The estimates indicate that there may be similar effects of earlier grades also in 6th vs 7th grade, increasing the probability of mental illness among girls. Although, the effects at this age are smaller and not statistically significant. When the effect is stratified like this, we also find a negative (protective) effect of earlier grading in 7th vs 8th grade on internalizing disorders among boys. Worth noting is that running multiple simultaneous tests like this increases the probability of erroneous inferences, thereby lone effects that differ from the main results should be interpreted with caution.

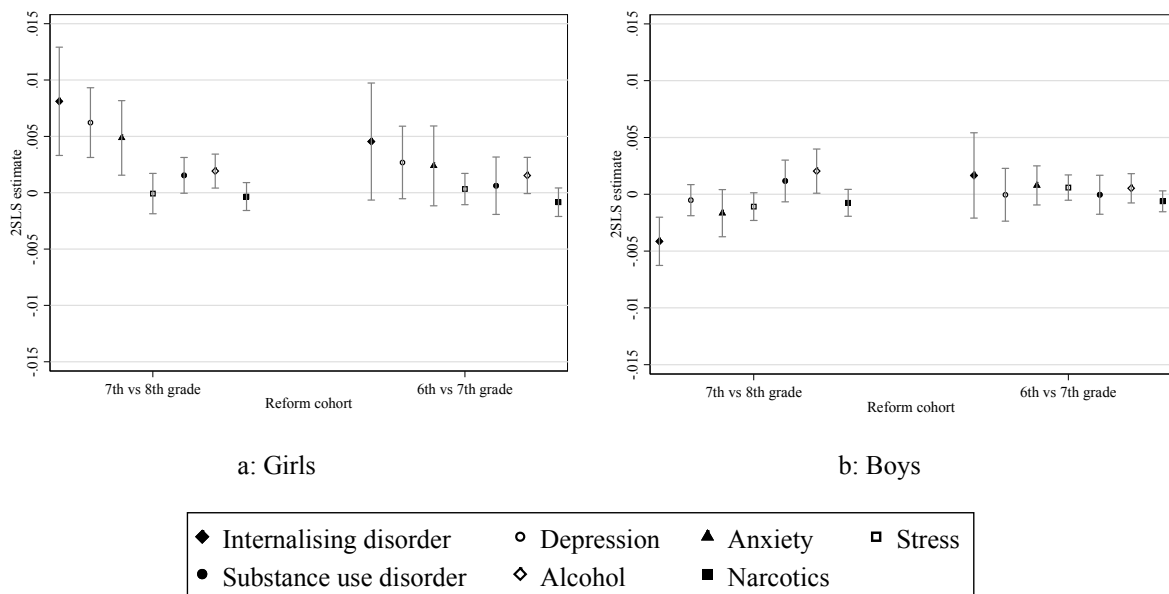


Figure 5: Sensitivity to Model Specification - Effect Heterogeneity by Age/School Year

Note: This figure shows the effect of earlier grades on the probability of mental illness in the year the child enters 9th grade in compulsory school. Spikes represent 95 % clustered along birth month and year confidence intervals. Each marker represents the effect estimate from separate difference-in-discontinuities regressions (ρ_{2SLS} , similarly as the main results in Table 2) but now treatment is separated in the reform cohorts, that is, the effect of earlier grades in 7th vs 8th grade, and the effect of grades in 6th vs 7th grade are investigated separately. The sample consists of all children born in Sweden six months before and after January 1st in 1999 (2000) for estimations on 7th vs 8th (6th vs 7th) grade, as well as the control cohorts born six months before and after January 1st in 1993, 1994, and 1998 (see data section for specific information about the sample).

5.4 Timing

In the analyses presented so far, we measure the impact of the reform on mental health in the year when children enter the final year of compulsory school (9th grade). However, the mental health response to earlier grading in schools may occur also at other points in time. Perhaps the anticipation of exposure to earlier grades induces such effects in prior years, or the introduction of earlier grades reduces mental illness in 8th grade related to pre-reform introduction of grades. We would also like to understand something about the persistence of the effects. Therefore, in this section, we investigate the effect of earlier grades on mental illness in the years the children enter 7th, 8th, and 9th grade, as well as the following year when a majority of the children start upper secondary school. This could reveal potential anticipation effects as well as persistence of the effects. To provide the best understanding of the mechanisms at play, we will perform this timing analysis in the reform cohorts separately (in 7th vs 8th and 6th vs 7th grades separately).

The results of earlier grading on mental illness by year, are shown in figure 6. These results support previous findings among boys, that effects of earlier grades on mental health are small. In fact, as shown in figure 6, if anything, the mental health response among boys to earlier grading appears protective, at least for internalizing disorders such as depression, anxiety and stress, but estimates are generally small and rarely significantly different from zero. Among girls, this timing analysis shows that the impact of grades in 7th instead of 8th grade on increased probability of diagnosis for all internalizing disorders, depression and anxiety in the year entering 9th grade starts to show effects already in previous years, suggesting that there is an immediate health response to earlier grades but that the main effect appears to be in the year entering the final year of compulsory school. The effect then appears to reduce in the year when they enter upper secondary school. Unfortunately, we cannot investigate even more long-term effects since the final year of our observation period is 2016. Moreover, the increased probability of alcohol-related diagnosis in 9th grade, among boys and girls, is preceded by a significantly lower probability of such diagnoses in the previous year, but similarly for these mental illness outcomes the effect reduces in the final follow-up year.

Effects of earlier grades in 6th instead of 7th grade are, in general, as the baseline result, smaller and not statistically significant. There is an increased probability of stress in the final follow-up year among girls, and a lower probability of alcohol-related diagnosis in the year

entering 8th grade among girls, but there are few consistencies in these results and considering multiple hypothesis testing problems we should interpret these findings with caution.

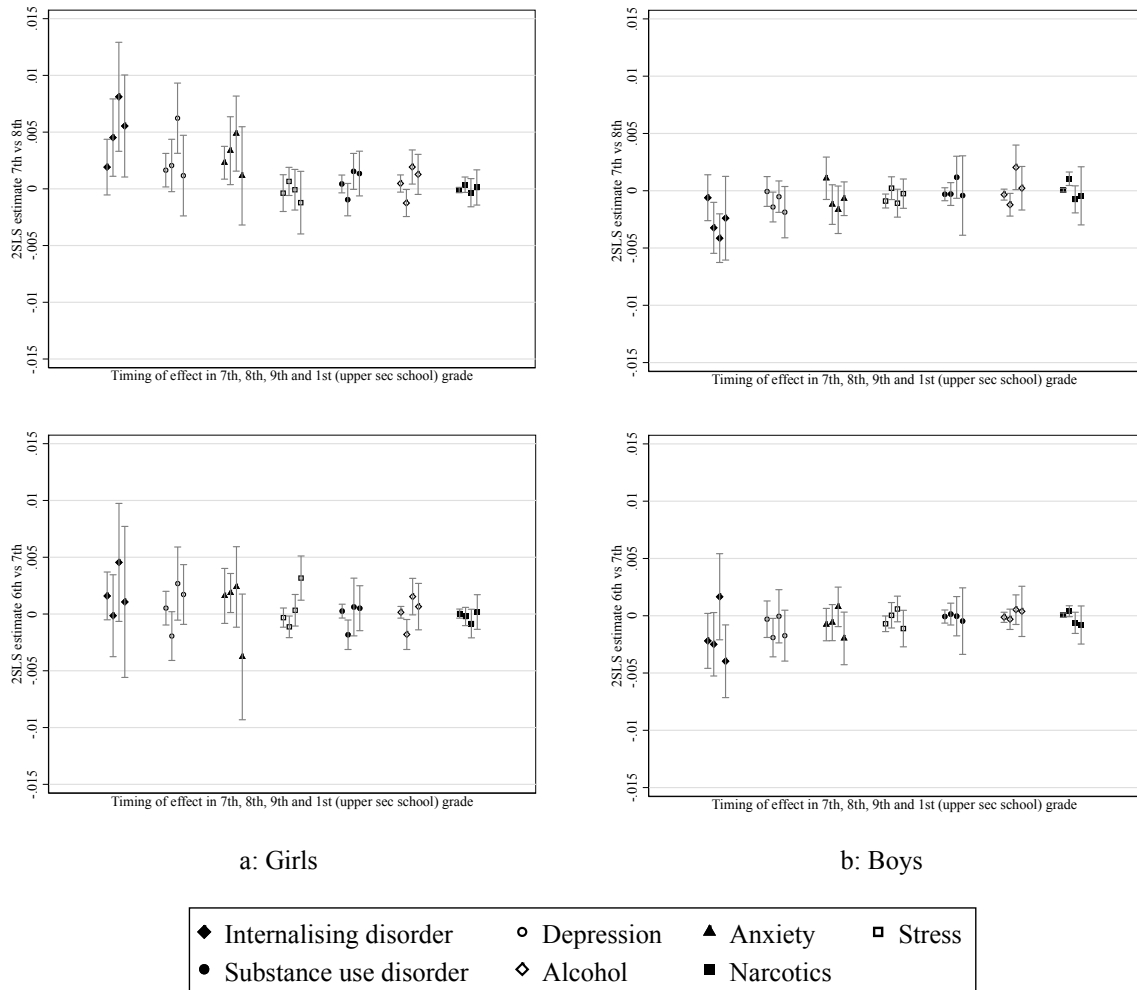


Figure 6: Timing of the effect

Note: This figure shows the timing of the effect of earlier grades on the probability of mental illness in the year the child enters 7th, 8th, 9th grade in compulsory school and the 1st year in Upper Secondary School. Each marker represents the effect estimate from separate difference-in-discontinuities regressions (p2SLS). Treatment is separated in the reform cohorts, that is, the effect of earlier grades in 7th vs 8th grade, and the effect of grades in 6th vs 7th grade are investigated separately. Spikes represent 95 % clustered along birth month and year confidence intervals. The sample consists of all children born in Sweden six months before and after January 1st in 1999 (2000) for estimations on 7th vs 8th (6th vs 7th) grade, as well as the control cohorts born six months before and after January 1st in 1993, 1994, and 1998 (see data section for specific information about the sample).

6. Robustness

6.1 Test for Manipulation

A general threat to identification in the regression discontinuity framework is manipulation of the threshold, in this case, if children (or their parents) manipulate the school cohort assignment to fit their preferred age of introduction to grading in school. It is unlikely that this type of manipulation would affect our results since changing school cohorts is quite

uncommon in Swedish schools, and the decision on the earlier grades reform was voted through in 2010, that is, when the relevant cohorts were already in compulsory school. To show that there is no manipulation, we examine the frequency of births by month of birth in the reform and control cohorts separately. As shown in figure 7, the trends in births, as well as the irregularity around the school cohort cutoff appear to be comparable between the reform and control cohorts.

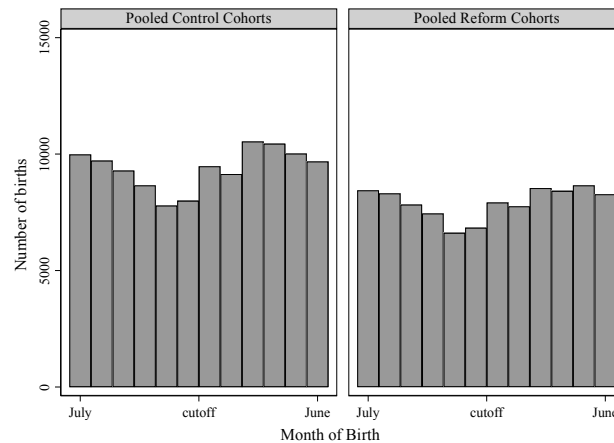


Figure 7: Frequency of Births

Note: This figure shows mean number of births by month of birth running from July to June for the reform cohorts, all children born in Sweden six months before and after January 1st in 1993, 1994, and 1998, and the control cohorts, all children born in Sweden six months before and after January 1st in 1999 and 2000 (see data section for specific information about the sample).

6.2 Covariate Balance

To further support our model, we show that month of birth is unrelated to several predetermined characteristics, such as migration background and parents' SES. If treatment is unrelated to these observed factors, it implies that treatment is balanced also in terms of unobserved factors. In figure 8, the mean probability of the selected covariates is plotted against month birth in discontinuity graphs, for the reform cohorts and control cohorts separately. In our Diff-in-Disc setup, rather than complete continuity at the threshold, as is required for standard RD setup, we only require that any discontinuity at the threshold is similar between the reform and control cohorts. That aligns with the parallel trends assumption discussed in the empirical strategy section. To complement the discontinuity graphs, we also include the covariates as dependent variables in Eq. 4, to investigate if earlier grades predict any of these factors (see table 4). The discontinuity graphs indicate that our treatment and control cohorts are balanced. The balancing test reveals a weak association between earlier grades and foreign-born parents which should be noted, but in general, this

test shows that treatment is independent of the predetermined factors, which implies that our model is a valid model to investigate the effect of earlier grades on mental health.

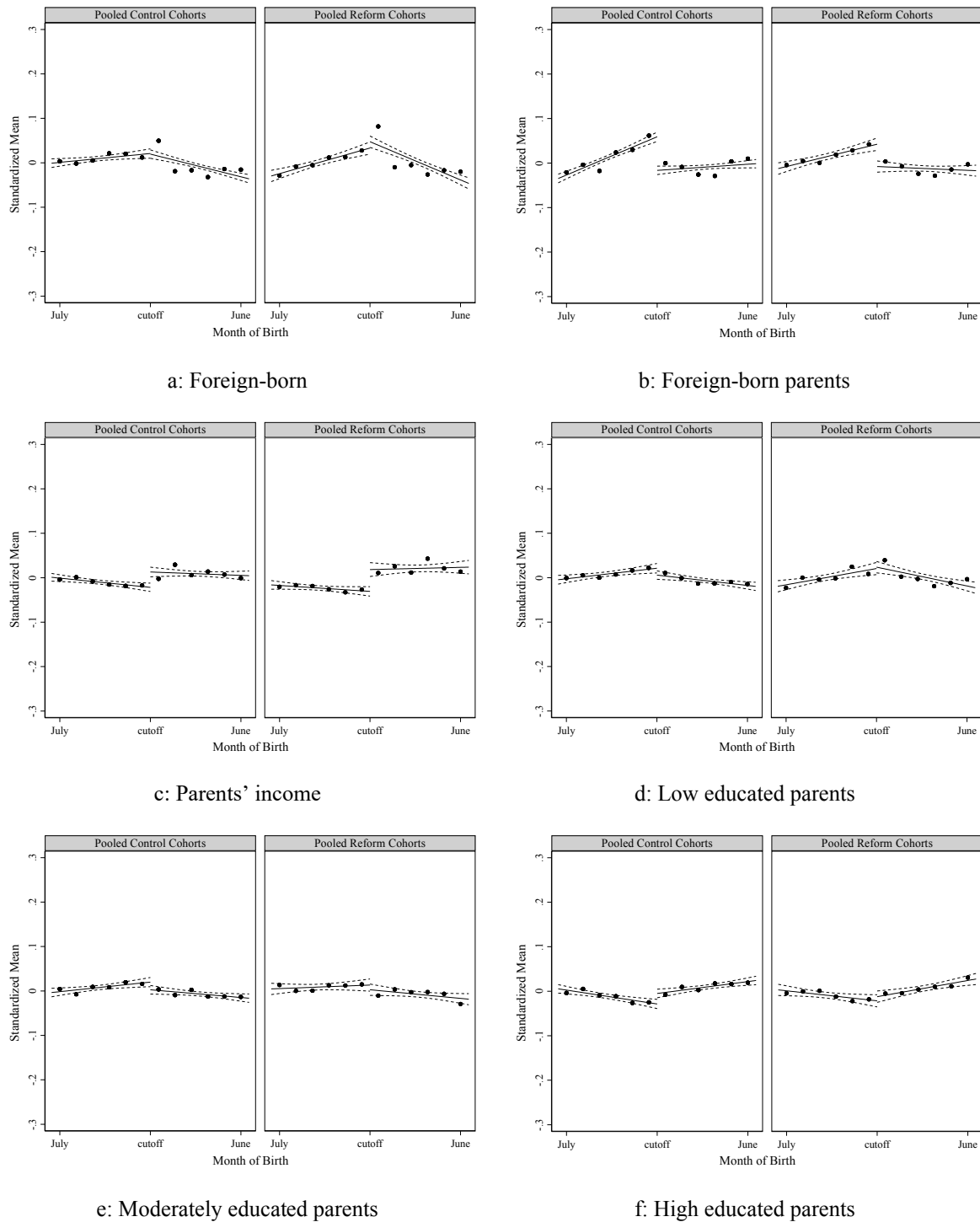


Figure 8: Discontinuity Graphs – Covariate Balance

Note: In this figure, the standardized mean of several predetermined characteristics and the fitted linear predictions of these characteristics (including 95 % confidence intervals) are plotted by month of birth (July to June). The plots are provided separately for the pooled control cohorts (those born around January 1st in years 1993, 1994 and 1998), and the pooled reform cohorts (those born around January 1st in 1999 and 2000).

Table 4: Covariate Balance, 2SLS

VARIABLES	(1) Foreign-born	(2) Foreign-born parents	(3) Low educated parents	(4) Moderately educated parent	(5) High educated parent	(6) Parents' income
Girls						
Earlier grades reform	0.00908 (0.00555)	0.0117* (0.00610)	0.00329 (0.00401)	0.0125 (0.00980)	-0.0158 (0.00976)	4,357 (5,762)
Constant	0.0960*** (0.00248)	0.127*** (0.00281)	0.0470*** (0.00185)	0.507*** (0.00454)	0.446*** (0.00450)	255,603*** (1,944)
<i>Observations</i>	254,901	254,901	254,901	254,901	254,901	241,393
Boys						
Earlier grades reform	0.00126 (0.00541)	0.00447 (0.00598)	0.00518 (0.00393)	-0.00425 (0.00966)	-0.000929 (0.00961)	5,452 (4,819)
Constant	0.0887*** (0.00240)	0.124*** (0.00281)	0.0445*** (0.00185)	0.515*** (0.00456)	0.440*** (0.00452)	253,822*** (1,914)
<i>Observations</i>	269,192	269,192	269,192	269,192	269,192	255,008

Note: This table presents the results for the impact of earlier grades (in 6th vs 7th and 7th vs 8th grade) on several predetermined characteristics. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (ρ_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample).

Cluster along birth month and year robust standard errors in the parentheses. ***p<0.01, **p<0.05, *p<0.1

6.3 Placebo-test

We now test the robustness of our results using placebo tests where we treat each separate control cohort as the reform cohort and measure the impact on mental illness (where we would expect zero effect). These tests show that there is no effect of “placebo exposure” to earlier grading on internalizing disorders, depression, and anxiety among girls (see figure 9). There is a positive relationship between placebo exposure in 1993 and substance use disorders (driven by alcohol-related disorders) among girls. For boys, there is a protective relationship with substance use- and alcohol-related disorders for a placebo reform affecting children born around January 1st in 1993 and a positive relationship on the same outcomes for a placebo reform affecting children born around January 1st in 1994. This placebo test supports our main results of an effect of earlier grades on internalizing disorders among girls, but, in line with previous results, implies that there is some uncertainty regarding the results on the different substance use disorders, especially among boys. This suggests that any results on these outcomes should be interpreted with caution.

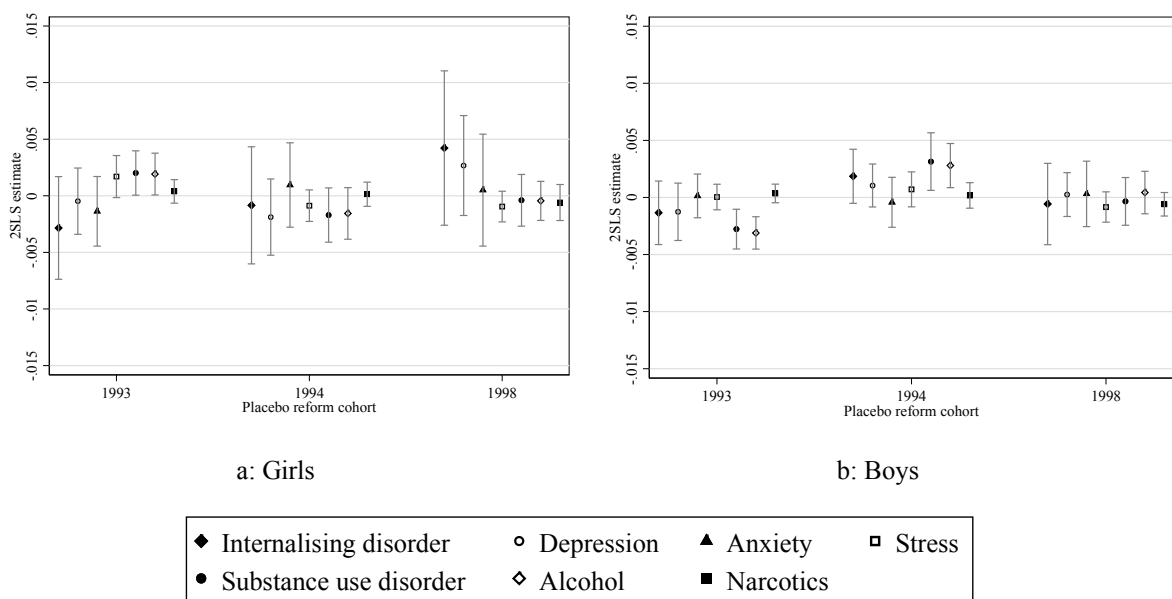


Figure 9: Placebo analysis

Note: This figure shows the effect of placebo-reform exposure to earlier grades on the probability of mental illness in the year the child enters 9th grade in compulsory school. Each marker represents the effect estimate from separate difference-in-discontinuities regressions (p_{2SLS} , similarly as the main results in Table 2). Spikes represent 95 % clustered along birth month and year confidence intervals. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, and 1998 (see data section for specific information about the sample).

7. Heterogeneity Analyses

Our results show that the impact of earlier grades on mental health differs between girls and boys. There may be heterogeneity also based on other background characteristics. In this section, we test if the effect of earlier grades on mental illness differs with parents' socioeconomic (income and education) and migration background.

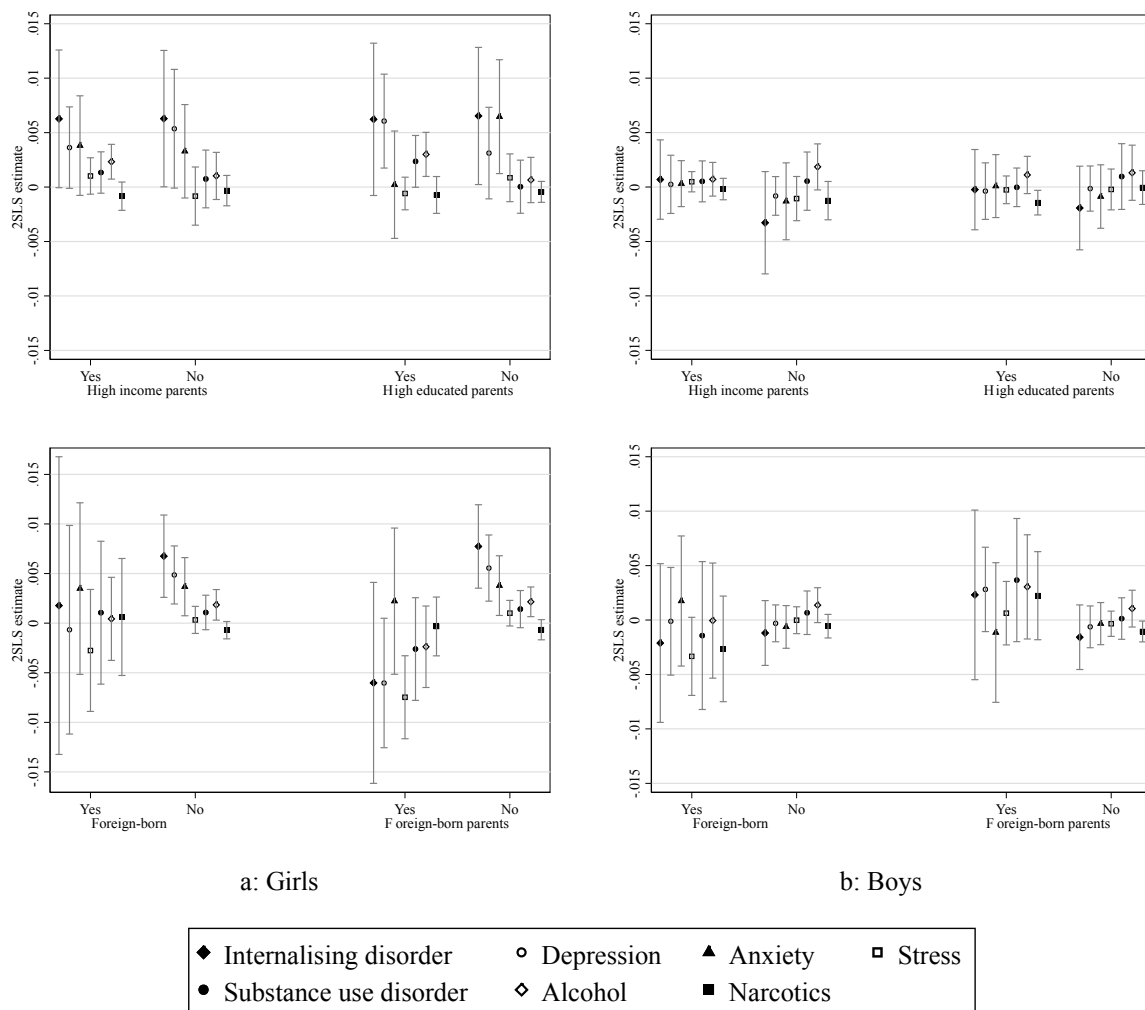


Figure 10: Heterogeneity analysis

Note: This figure shows heterogeneity in the effect of earlier grades on the probability of mental illness in the year the child enters 9th grade in compulsory school. Each marker represents the effect estimate from separate difference-in-discontinuities regressions (ρ_{2SLS} , similarly as the main results in Table 2) but in groups stratified by above/below parents' median income, at least one higher educated parent or not, foreign-born or not and foreign-born parents or not. Spikes represent 95 % clustered along birth month and year confidence intervals. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample).

The results indicate that there may be heterogeneity in the effect of earlier grading on mental illness among girls (see figure 10). Among boys, the results reflect the baseline results. No effects of earlier grading on mental illness in any of the stratified groups can be seen. The effect of earlier grades on increased probability of internalizing disorders such as depression

and anxiety among girls does not appear to differ much with parents' socioeconomic background, except that there is a stronger effect on depression for girls with at least one high-educated parent and a stronger effect on anxiety for girls with no high-educated parent. The results moreover show that the effects of earlier grading on any internalizing disorder, depression, and anxiety diagnosis, are significantly different from zero only among girls who are born in Sweden and have Swedish-born parents when stratified by migration background. If anything, the results indicate that having foreign-born parents is protective for depression and stress-related diagnoses among girls, but statistically significant only for stress-related diagnoses. The results moreover indicate that the effect on increased probability of alcohol-related disorders appears to be strongest among girls with high-income and high-educated parents compared to girls with low SES parents, and among girls with no foreign background compared to girls with foreign backgrounds. Though, as mentioned before, the results on all substance use disorders are unreliable.

8. Discussion

Our results show that being exposed to earlier grading in compulsory school in Sweden increases the probability of being diagnosed with internalizing disorders such as anxiety and depression in the end of compulsory school, but only among girls. The absolute effect estimates are modest, 0.63, 0.44, and 0.36 percentage points increase in the probabilities for any internalizing disorder, depression, or anxiety diagnosis, respectively. However, given that diagnosis for these disorders is uncommon at this age, around one and a half percent for any internalizing disorder and just below one percent for depression and anxiety, the results imply that earlier grades increase the probability of diagnosis by roughly 50-70 percent depending on the outcome.

In general, our results also indicate a positive relationship between earlier grades and alcohol-related diagnosis, but the absolute probabilities for all the substance use disorder outcomes (alcohol and narcotics) are very low and appear unstable to modelling choices. These inconsistencies suggest that we should interpret the findings on substance use disorders carefully, but they also indicate interesting relationships that should be further investigated, perhaps using more long-term follow-up data, or investigating measures that are more sensitive among a young population.

We find that it is mainly the effect of one year earlier grading in grade 7, and not in grade 6, which increases the probability of diagnosis. Entering grade 7 often means that you must change school. Thus, it may be the combination of starting a new school, new classmates, and mentors, and at the same time being introduced to grades in school for the first time, that contributes to the increasing mental illness. Introducing grades already in grade 6 may be considered a relatively smaller adjustment which could explain the heterogeneity in the effects. Another possible explanation is that grading in grade 6 is based on the learning objectives and knowledge requirements set for the grade, while grading in grade 7 is based on the learning objectives and knowledge requirements for grade 9, even if consideration should be taken to the relative grade level. Beliefs about one's own abilities in relation to the set knowledge requirements could explain that we find an effect of earlier grading in grade 7, but not in grade 6, on mental illness. It should not be concluded that there are no adverse effects of introducing grades in grade 6. Since we measure the impact of earlier grades in 6th vs 7th grade and in 7th vs 8th grade, an effect in grade 7 and no effect in grade 6 just means that there is no additional effect of introducing grades in grade 6. Children who receive grades for the first time in grade 6 and grade 7 are similarly worse off compared to those who receive grades for the first time in grade 8. Moreover, it is worth noting that there are some indications in the results of an additional effect of earlier grades also in grade 6, even though this effect was not statistically significant.

We find clear differences across the sexes, where earlier grades increase the probability of internalizing disorders among girls, but not boys. Högberg et al. (2021) also find that the effect of earlier grades is stronger among girls, but only for the impact on life satisfaction. We are not aware of any other studies examining the extensive margin of grading effects on health, but the previous work by Klapp (2015) and Sjögren (2010) also suggests that the effects of earlier grades on school performance and attainment are stronger among girls. Studying effects of grading bias on mental health, Linder et al. (2023) show that being exposed to lenient grading has protective effects on young women's mental health.

Why then, is the mental health of girls more responsive to grades than boys? It has been suggested that during adolescence the importance of performing well in school is larger among girls than boys, and therefore they experience more stress and pressure related to school performance and the assessment of such performance (West and Sweeting 2003). Previous research also finds that women value grades more than men and argue that this is an

explanation for the gender gap in STEM-tracks; because STEM-classes generally are lower graded compared to non-STEM classes they are perceived as lower benefit among female students (Ahn, Arcidiacono et al. 2019). Girls generally have much higher prevalence of internalizing mental illness, both in terms of received diagnoses and pharmaceutical prescriptions (The Swedish National Board on Health and Welfare 2017), as well as in terms of self-reported mental illness (The Public Health Agency of Sweden 2018). This may, of course, be an indication that females actually experience different levels of depression and anxiety than males, but the gap between the genders may also, at least to some extent, be explained by that females and males seek mental health care to a varying extent, and at different symptom levels, in which case a diagnosis may reflect different levels of severity of the mental illness (Kovess-Masfety et al. 2014). Similarly, self-reported mental problems are also dependent on the actual reporting of such problems. Differences in self-reported health may reflect differences in the tendency of reporting problems between groups. Similarly, if a diagnosis among girls captures something different to that a diagnosis among boys captures, this may explain why we find effects of earlier grading among girls but not among boys. The heterogeneity in how girls and boys respond to grades, and how this may impact the existing mental health gap between females/males, need to be further investigated.

In general, the research on feedback and assessment of students indicates that effects differ based on other background factors such as parents' education but also cognitive ability, see for example Sjögren (2010), Dee and Jacob (2011), Klapp (2015), Whitney and Candelaria (2017). Our results indicate that the effects of earlier grading on mental illness (internalizing diagnoses) differs only slightly by socioeconomic background, but there appears to be heterogeneity related to migration background. The effects of earlier grades on increased probabilities of any internalizing disorder, depression and anxiety are driven by girls that have no migration background (Swedish born to Swedish parents). Moreover, the results indicate a protective effect of earlier grades on stress-related diagnoses among girls with foreign-born parents. This sole effect does not constitute basis for conclusions regarding how migration background affects the mental health response to grades among children in school, but we find similar patterns also for other outcomes, depression, and alcohol-related diagnoses. Even though these effects are non-significant, they indicate that heterogeneity related to migration background in this relationship may be of importance, which needs to be further investigated.

Our study contributes to the literature on performance feedback and human capital development. The results of our study are based on a robust identification strategy subject to several robustness and sensitivity checks such as the inclusion of additional control cohorts, a placebo test, the inclusion of higher order polynomials, alternative bandwidths, and balance across predetermined characteristics, which supports the validity of our empirical strategy. A slight increasing trend is noted in the pre-treatment period for internalizing disorders (see figure 3) which suggests that conclusions regarding this outcome should be made with caution, but in general, the effects of earlier grading on increased depression and anxiety among girls are very stable across specifications and robust to several tests of the underlying model assumptions.

The data is based on administrative records that have low measurement errors and allow for detailed heterogeneity analysis. There are however limits to the conclusions we can draw. We have only considered mental illness measured by diagnoses within inpatient and specialized outpatient care in Sweden. More common and sensitive indicators of mental health problems among this age group, for example, prescription of antidepressants or anxiolytics, may yield different and more precise results than that we have found. We are also not able to consider longer-term outcomes due to our sample period. The results show that the effect starts to build up immediately after the grading in grade 7, although the main effect seems to appear only in the year when the child enters in grade 9. Our findings then show this effect diminishes or disappears in the following year. More research is needed to understand how earlier grades affect children and young people's human capital development in the longer term, both by examining the health impacts of the reform in the years that follow, and to examine possible mechanisms such as academic performance in upper secondary school.

9. Conclusion

In this paper, we show that introducing earlier grades in compulsory schools in Sweden increased the probability of being diagnosed for internalizing disorders such as depression and anxiety, among girls exposed to the reform. Our findings imply that part of the increasing mental illness trend among girls and young women may in part be attributed to changes in school processes related to the shift towards increased educational accountability. This reveals an important health mechanism that may be of importance for the substantial gaps observed in internalizing mental health that occurs between girls and boys in school age and persists throughout life between women and men.

References

- Ahn, T., P. Arcidiacono, A. Hopson and J. R. Thomas (2019). "Equilibrium Grade Inflation with Implications for Female Interest in STEM Majors." National Bureau of Economic Research Working Paper Series No. 26556.
- Brismark, A., A. Hullung, J. Samuelsson and M. Stolare (2016). "Det är mer på allvar nu" Föräldrars erfarenheter av betyg i årskurs 6. The Swedish National Agency for Education Stockholm. Access date: 2023-02-27. Available from: <https://www.skolverket.se/publikationer?id=3736>
- Currie, J. and M. Stabile (2006). "Child mental health and human capital accumulation: The case of ADHD." Journal of Health Economics 25(6): 1094-1118.
- Currie, J. and M. Stabile (2007). "Mental health in childhood and human capital." NBER Working Paper Series.
- Currie, J., M. Stabile, P. Manivong and L. L. Roos (2010). "Child Health and Young Adult Outcomes." The Journal of Human Resources 45(3): 517-548.
- Collins, M. and J. Lundstedt (2021). "The Long-term Consequences of More Informative Grading." LUND PAPERS IN ECONOMIC DEMOGRAPHY 2021(1)
- Dee, T. S. and B. Jacob (2011). "The impact of no Child Left Behind on student achievement." Journal of Policy Analysis and Management 30(3): 418-446.
- Ding, W., S. F. Lehrer, J. N. Rosenquist and J. Audrain-McGovern (2009). "The impact of poor health on academic performance: New evidence using genetic markers." Journal of Health Economics 28(3): 578-597.
- Figlio, D. and S. Loeb (2011). Chapter 8 - School Accountability. Handbook of the Economics of Education. E. A. Hanushek, S. Machin and L. Woessmann, Elsevier. 3: 383-421.
- Fletcher, J. M. (2010). "Adolescent depression and educational attainment: results using sibling fixed effects." Health Econ 19(7): 855-871.
- Fredriksson, P. and B. Öckert (2014). "Life-cycle Effects of Age at School Start." The Economic Journal 124(579): 977-1004.
- Grembi, V., T. Nannicini and U. Troiano (2016). "Do Fiscal Rules Matter?" American Economic Journal: Applied Economics 8(3): 1-30.
- Gustafsson, J.-E., M. Allodi, A. Westling, B. Åkerman, C. Eriksson, L. Eriksson, S. Fischbein, M. Granlund, P. Gustafsson, S. Ljungdahl, T. Ogden and R. S. Persson (2010). School, learning and mental health : a systematic review. Stockholm, Kungl. Vetenskapsakademien. Hälsoutskottet
- Heckman, J. J. (2012). "The developmental origins of health." Health Econ, 21(1), 24-29.
- Högberg, B. (2021). "Educational stressors and secular trends in school stress and mental health problems in adolescents." Social Science & Medicine 270: 113616.
- Högberg, B., J. Lindgren, K. Johansson, M. Strandh and S. Petersen (2021). "Consequences of school grading systems on adolescent health: evidence from a Swedish school reform." Journal of Education Policy 36(1): 84-106.

- Klapp, A. (2015). "Does grading affect educational attainment? A longitudinal study." Assessment in Education: Principles, Policy & Practice 22(3): 302-323.
- Klapp, A., C. Cliffordson and J.-E. Gustafsson (2016). "The effect of being graded on later achievement: evidence from 13-year olds in Swedish compulsory school." Educational Psychology 36(10): 1771-1789.
- Kovess-Masfety, V., A. Boyd, S. van de Velde, R. de Graaf, G. Vilagut, J. M. Haro, S. Florescu, S. Neill, L. Weinberg, J. Alonso and E. U. W. M. H. i. the (2014). "Are there gender differences in service use for mental disorders across countries in the European Union? Results from the EU-World Mental Health survey." Journal of Epidemiology and Community Health 68(7): 649.
- Linder, A., M. Nordin, U.-G. Gerdtham and G. Heckley (2023). "Grading bias and young adult mental health." Health Economics 32(3): 675-696.
- Lingard, B., W. Martino and G. Rezai-Rashti (2013). "Testing regimes, accountabilities and education policy: commensurate global and national developments." Journal of Education Policy 28(5): 539-556.
- Lundahl, C., M. Hultén and S. Tveit (2016). Betygssystem i internationell belysning. Stockholm, The Swedish National Agency for Education. Access date: 2023-02-27. Available from: <https://www.skolverket.se/publikationsserier/ovrigt-material/2017/betygssystem-i-internationell-belysning>
- Lundahl, C., M. Hultén and S. Tveit (2017). "The power of teacher-assigned grades in outcome-based education." Nordic Journal of Studies in Educational Policy 3(1): 56-66.
- Potrebny, T., N. Wiium and M. M.-I. Lundegård (2017). "Temporal trends in adolescents' self-reported psychosomatic health complaints from 1980-2016: A systematic review and meta-analysis." PLOS ONE 12(11): e0188374.
- Sjögren, A. (2010). Graded children: Evidence of longrun consequences of school grades from a nationwide reform. IFAU working paper 2010:7.
- The Swedish Government (2009). Regeringens proposition 2009/10:219. Betyg från årskurs 6 i grundskolan Utbildningsdepartementet. Stockholm. Access date: 2023-02-27. Available from: <https://www.regeringen.se/rattsliga-dokument/proposition/2010/09/prop.-200910219>
- The Public Health Agency of Sweden (2018). Varför har den psykiska ohälsan ökat bland barn och unga i Sverige? The Public Health Agency of Sweden. Stockholm. Access date: 2023-02-27. Available from: <https://www.folkhalsomyndigheten.se/contentassets/628f1bfc932b474f9503cc6f8e29fd45/var-for-psykiska-ohalsan-okat-barn-unga-18023-2-webb-rapport.pdf>
- The Swedish National Agency for Education (2011). Läroplan för grundskolan, förskoleklassen och fritidshemmet. Stockholm, The Swedish National Agency for Education. Access date: 2023-02-27. Available from: <https://www.skolverket.se/download/18.35e3960816b708a596c3965/156767>
- The Swedish National Agency for Education (2017). Utvärdering av betyg från årskurs 6. Stockholm, The Swedish National Agency for Education. Rapport 451. Access date: 2023-02-27. Available from: <https://www.skolverket.se/download/18.6bfaca41169863e6a65c7da/1553967200750/pdf3735.pdf>

The Swedish National Board on Health and Welfare (2017). Utvecklingen av psykisk ohälsa bland barn och unga vuxna - Till och med 2016. The Swedish National Board on Health and Welfare Stockholm. Access date: 2023-02-27. Available from:

<https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2017-12-29.pdf>

Vlachos, J. (2011). Betygets värde. En analys av hur konkurrens påverkar betygssättningen vid svenska skolor. Stockholm: Konkurrensverket, Konkurrensverkets uppdragsforskningsrapport 2010:6.

West, P. and H. Sweeting (2003). "Fifteen, female and stressed: changing patterns of psychological distress over time." Journal of Child Psychology and Psychiatry 44(3): 399-411.

Whitney, C. R. and C. A. Candelaria (2017). "The Effects of No Child Left Behind on Children's Socioemotional Outcomes." AERA Open 3(3): 2332858417726324.

Wikström, C. and M. Wikström (2005). "Grade inflation and school competition: an empirical analysis based on the Swedish upper secondary schools." Economics of Education Review 24(3): 309-322.

Appendix

Table A1: First stage

	(1)
VARIABLES	Earlier grades reform
Born January-June*reform (earlier grades reform instrument)	0.967***
	(0.000496)
Constant	0.0100***
	(0.00284)
Observations	524,093
R-squared	0.921

Note: This table presents the results for the first stage of the instrument, that is the effect of being born in January to June in a reform year on actual exposure to the reform which is calculated based on coherence between birth cohort and school cohort, school cohort derived from the year of taking the national exams in 9th grade. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample).

Cluster along birth month and year robust standard errors in the parentheses. ***p<0.01, **p<0.05, *p<0.1

Table A2: Robust standard errors - The Effect of Earlier Grading on Mental Illness at the End of Compulsory School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Any internalizing diagnosis	Depression	Anxiety	Stress	Any substance use diagnosis	Alcohol	Narcotics
Full sample							
Earlier grades reform	0.00244*	0.00202**	0.00151	-6.05e-05	0.000796	0.00148**	-0.000643
	(0.00145)	(0.000999)	(0.00105)	(0.000527)	(0.000712)	(0.000590)	(0.000423)
Constant	0.00944***	0.00414***	0.00552***	0.00127***	0.00288***	0.00227***	0.000718***
Observations: 524,093	(0.000633)	(0.000420)	(0.000448)	(0.000249)	(0.000396)	(0.000359)	(0.000186)
Girls							
Earlier grades reform	0.00629**	0.00441**	0.00359**	0.000139	0.00107	0.00172*	-0.000596
	(0.00248)	(0.00176)	(0.00181)	(0.000903)	(0.00107)	(0.000897)	(0.000623)
Constant	0.0132***	0.00627***	0.00784***	0.00180***	0.00379***	0.00333***	0.000646**
Observations: 254,901	(0.00106)	(0.000724)	(0.000753)	(0.000421)	(0.000602)	(0.000556)	(0.000267)
Boys							
Earlier grades reform	-0.00124	-0.000272	-0.000467	-0.000248	0.000542	0.00127	-0.000696
	(0.00154)	(0.000988)	(0.00111)	(0.000562)	(0.000943)	(0.000774)	(0.000575)
Constant	0.00599***	0.00220***	0.00339***	0.000801***	0.00201***	0.00128***	0.000780***
Observations: 269,192	(0.000704)	(0.000436)	(0.000495)	(0.000270)	(0.000518)	(0.000457)	(0.000260)

Note: This table presents the results for the impact of earlier grades (in 6th vs 7th and 7th vs 8th grade) on the probability of mental illness in the year the child enters 9th grade in compulsory school. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (ρ_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample). Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table A3: Including polynomials - The Effect of Earlier Grading on Mental Illness at the End of Compulsory School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Any internalizing diagnosis	Depression	Anxiety	Stress	Any substance use diagnosis	Alcohol	Narcotics
Full sample							
Earlier grades reform (incl 2 nd)	0.00295*	0.00167	0.00298***	0.000303	0.00278***	0.00328***	-0.000486
	(0.00163)	(0.00139)	(0.000985)	(0.000719)	(0.000687)	(0.000658)	(0.000407)
Earlier grades reform (incl 2 nd and 3 rd)	0.00335	0.00566***	0.000865	0.000809	0.00172*	0.00310***	-0.00117*
Observations: 524,093	(0.00210)	(0.00199)	(0.00124)	(0.000970)	(0.000969)	(0.000940)	(0.000628)
Girls							
Earlier grades reform (incl 2 nd)	0.00610***	0.00311	0.00574***	0.000283	0.00477***	0.00468***	0.000120
	(0.00225)	(0.00210)	(0.00161)	(0.000920)	(0.00124)	(0.000978)	(0.000766)
Earlier grades reform (incl 2 nd and 3 rd)	0.00685**	0.00912***	0.00456*	0.000694	0.00321*	0.00457***	-0.00138
Observations: 254,901	(0.00321)	(0.00311)	(0.00254)	(0.00133)	(0.00174)	(0.00133)	(0.00120)
Boys							
Earlier grades reform (incl 2 nd)	0.000278	0.000441	0.000481	0.000365	0.000844	0.00193	-0.00107*
	(0.00197)	(0.00121)	(0.00132)	(0.000866)	(0.00172)	(0.00143)	(0.000642)
Earlier grades reform (incl 2 nd and 3 rd)	0.000627	0.00268	-0.00245	0.00114	0.000330	0.00176	-0.000973
Observations: 269,192	(0.00293)	(0.00180)	(0.00176)	(0.00121)	(0.00265)	(0.00220)	(0.000975)

Note: This table presents the results for the impact of earlier grades (in 6th vs 7th and 7th vs 8th grade) on the probability of mental illness in the year the child enters 9th grade in compulsory school. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (ρ_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The regression includes the 2nd and 3rd order polynomials of the running variable (re-centered birth month). The sample consists of all children born in Sweden six months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample). Cluster along birth month and year robust standard errors in the parentheses. ***p<0.01, **p<0.05, *p<0.1

Table A4: Alternative bandwidth – The Effect of Earlier Grading on Mental Illness at the End of Compulsory School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Any internalizing diagnosis	Depression	Anxiety	Stress	Any substance use diagnosis	Alcohol	Narcotics
Full sample							
Earlier grades reform	0.00269**	0.00299**	0.00160*	0.000551	0.00206***	0.00272***	-0.000653*
	(0.00137)	(0.00123)	(0.000833)	(0.000603)	(0.000555)	(0.000559)	(0.000368)
Constant	0.00944***	0.00423***	0.00529***	0.00129***	0.00327***	0.00262***	0.000837***
Observations: 524,093	(0.000825)	(0.000449)	(0.000533)	(0.000174)	(0.000287)	(0.000242)	(0.000161)
Girls							
Earlier grades reform	0.00507**	0.00484**	0.00390**	0.000380	0.00324***	0.00372***	-0.000541
	(0.00214)	(0.00196)	(0.00160)	(0.000809)	(0.000997)	(0.000716)	(0.000725)
Constant	0.0140***	0.00668***	0.00820***	0.00183***	0.00436***	0.00367***	0.000980***
Observations: 254,901	(0.00140)	(0.000957)	(0.000812)	(0.000247)	(0.000412)	(0.000275)	(0.000250)
Boys							
Earlier grades reform	0.000817	0.00140	-0.000401	0.000784	0.000941	0.00178	-0.000764
	(0.00147)	(0.00118)	(0.000995)	(0.000737)	(0.00158)	(0.00131)	(0.000657)
Constant	0.00545***	0.00204***	0.00263***	0.000844**	0.00220***	0.00160***	0.000682***
Observations: 269,192	(0.000647)	(0.000366)	(0.000495)	(0.000347)	(0.000628)	(0.000562)	(0.000210)

Note: This table presents the results for the impact of earlier grades (in 6th vs 7th and 7th vs 8th grade) on the probability of mental illness in the year the child enters 9th grade in compulsory school. The results in each column and panel are from separate difference-in-discontinuities regressions, the effect estimates (ρ_{2SLS}) thus represents the LATE of exposure to one year earlier grading. The sample consists of all children born in Sweden three months before and after January 1st in 1993, 1994, 1998, 1999 and 2000 (see data section for specific information about the sample). Cluster along birth month and year robust standard errors in the parentheses. ***p<0.01, **p<0.05, *p<0.1