Family Spillover Effects of Marginal Diagnoses: The Case of ADHD*

Petra Persson† Xinyao Qiu‡ Maya Rossin-Slater§

January 6, 2021

Abstract

The health care system commonly relies on information about family medical history in the allocation of screenings and in diagnostic processes. At the same time, an emerging literature documents that treatment for "marginally diagnosed" patients often has minimal impacts. This paper shows that reliance on information about relatives' health can perpetuate marginal diagnoses across family members, thereby raising caseloads and health care costs, but without improving patient well-being. We study Attention Deficit Hyperactivity Disorder (ADHD), the most common childhood mental health condition, and document that the younger siblings and cousins of marginally diagnosed children are also more likely to be diagnosed with and treated for ADHD. Moreover, we find that the younger relatives of marginally diagnosed children have no better adult human capital and economic outcomes than the younger relatives of those who are less likely to be diagnosed. Our analysis points to a simple adjustment to physician protocol that can mitigate these marginal diagnosis spillovers.

JEL classification: I14, I18, J13

Keywords: ADHD, targeting, marginal diagnosis, mental health, family spillovers

*We thank Amitabh Chandra, Liran Einav, Ben Handel, Andrew Hertzberg, Jon Kolstad, Adriana Lleras-Muney, Nolan Miller, Ziad Obermeyer, and seminar participants at Stanford University, the Electronic Health Economics Colloquium, American University, UC Berkeley, Philadelphia Federal Reserve, McGill University, Penn State, and the University of Copenhagen for helpful comments. We thank Katja Hoffman for excellent research assistance.

†Department of Economics, Stanford University; NBER; CEPR; Research Institute for Industrial Economics, Stockholm. Email: perssonp@stanford.edu.

‡Department of Economics, Stanford University. Email: xyqiu@stanford.edu.

§Department of Medicine, Stanford University School of Medicine; NBER; IZA. Email: mrossin@stanford.edu.

1 Introduction

For hereditary diseases, an individual's diagnosis contains information about the risk of the condition for their family members. Thus, the health care system often relies on family medical history in the allocation of screenings and in diagnostic processes. For example, if a woman is found to carry particular mutations of a breast cancer gene (BRCA), then her close female family members are typically referred to genetic screening for BRCA.[1] The benefits of such "hereditary tagging" are clear: Screening the relatives of previously diagnosed patients allows the health care system to target scarce screening resources toward *ex ante* high-risk individuals, which reduces the social cost of identifying patients who need medical treatment in the population.

At the same time, an emerging literature documents that health care treatment for *marginally* diagnosed patients often has minimal or even negative impacts on patient health and well-being (Alalouf et al., 2019; Bos et al., 2020; Cuddy and Currie, 2020), while a closely related literature argues that a variety of conditions are frequently misdiagnosed (see, e.g., Mullainathan and Obermeyer, 2017; Obermeyer et al., 2019).[2] In this paper, we document that the use of family medical history can perpetuate low (or even negative) value marginal diagnoses across family members, thereby raising caseloads and health care costs, but without improving patient well-being. Our analysis also points to a simple adjustment to physician protocol that can mitigate these marginal diagnosis spillovers.

Specifically, we study this issue in the context of Attention Deficit Hyperactivity Disorder (ADHD), the most common mental health condition among children, affecting nearly ten percent of children in the United States (Danielson et al., 2018) and seven percent of children worldwide (Thomas et al., 2015). ADHD is characterized by a range of symptoms, including having trouble paying attention, staying organized, and remembering details. While the full

[1] See `https://www.cdc.gov/genomics/disease/breast_ovarian_cancer/testing.htm` for more details about BRCA gene testing.

[2] Mullainathan and Obermeyer (2017) and Obermeyer et al. (2019) highlight the role of machine learning algorithms in propagating misdiagnoses, biases, and the mis-allocation of health care treatment. Additionally, there exist studies on over-diagnoses of breast cancer (Brewer et al., 2007; Bond et al., 2013; Ong and Mandl, 2015; Einav et al., 2019) and pneumonia (Chan et al., 2019). One interpretation of low-value "marginal" diagnoses is that they are erroneous. Another interpretation is that the scientifically agreed-upon threshold for diagnosing a condition is "too low"; that is, even if a particular marginal diagnosis is not erroneous *per se*, "the cure" that comes with a diagnosis is, from the patient's perspective, no better than "the disease."

set of causes is unknown, the etiology of ADHD has a strong genetic component (Levy et al., 1997; Thapar and Cooper, 2016; Miller et al., 2019).[3]

Our empirical design exploits a well-documented fact about ADHD: Children who are younger for their grade level are on the margin more likely to be diagnosed and treated than their older classmates.[4] This diagnosis gap is typically interpreted as reflecting differences in maturity between children who are almost one year apart in age—children who are youngest in the classroom naturally have more difficulties paying attention and sitting still than their older classmates. If one does not account for differences in children's relative age for grade, then one may misinterpret these differences in maturity as differences in ADHD prevalence. We use population-level Swedish administrative data on children born in 1990–1996, and start by confirming this previously documented phenomenon in our data and sample. We find that December-born children—who are born just before the Swedish school entry cutoff of January 1st and are therefore youngest in their grade—are 30.8 percent more likely to be diagnosed with ADHD and 29.5 percent more likely to be treated with ADHD medication than their January-born peers.

We then take advantage of our ability to link family members in the data to document two new facts about marginal diagnoses of ADHD. First, we show that the age-for-grade-induced marginal diagnoses propagate to younger family members. The younger siblings and cousins of children who are born just before the school entry cutoff are 12.0 and 13.0 percent more likely to be diagnosed with ADHD, and 9.8 and 9.1 percent more likely to be treated with ADHD drugs, respectively, than the younger siblings and cousins of children born shortly after the cutoff. Importantly, these discontinuities exist *conditional on the younger siblings' and cousins' own relative age for grade*. In fact, the combined magnitudes of the estimated spillover effects on younger siblings and cousins amount to more than half of the younger children's own relative age effects on ADHD diagnoses and drug treatment.

Second, we examine the longer-term human capital and economic outcomes of the younger children. Conditional on their own relative age for grade, younger siblings of children who are

[3] Also see, e.g., Faraone et al. (1992); Barkley (2006); Tarver et al. (2014).

[4] See, for example: Elder, 2010; Evans et al., 2010; Dalsgaard et al., 2012; Morrow et al., 2012; Zoëga et al., 2012; Halldner et al., 2014; Krabbe et al., 2014; Pottegård et al., 2014; Chen et al., 2016; Schwandt and Wuppermann, 2016; Layton et al., 2018; Whitely et al., 2018; Root et al., 2019; Furzer et al., 2020; Furzer, 2020.

born before the cutoff—i.e., those who are disproportionately more likely to be diagnosed with and treated for ADHD—have 1.6 percent *lower* earnings at ages 23–28 and are 2.1 percent *less* likely to have ever enrolled in college than younger siblings of children born after the cutoff.[5] Younger cousins of children who are born before the cutoff have a 1.0 percent lower grade point average (GPA) in high school and are 2.6 percent less likely to have ever enrolled in college than the younger cousins of children born after the cutoff. The spillover effect on cousins' earnings is of the same sign as the one for siblings, but smaller in magnitude (and statistically insignificant). In sum, our results on educational and economic outcomes suggest that the younger siblings and cousins of marginally diagnosed children are no better off in the long-term; if anything, there could be long-term economic costs associated with marginal ADHD diagnosis spillovers.[6]

The existence of the ADHD diagnosis spillovers suggests that information about the older child's diagnosis is used in the process through which the younger child's ADHD diagnosis comes about. As described in Section 2, the process of obtaining an ADHD diagnosis involves two key steps: First, one must get a referral for an ADHD evaluation, and second, a physician performs an ADHD screening. To understand how the ADHD spillovers arise, we investigate the roles of the parties involved in each step.

"Hereditary tagging" in the referral stage would imply that an ADHD diagnosis of one child raises the likelihood that a younger family member receives a referral for an ADHD screening. Indeed, the referral stage is generally where such tagging can help to target scarce screening resources toward *ex ante* high-risk individuals. Schools and parents play important roles in requesting ADHD screenings, and we present evidence consistent with both entities responding to an older child's relative-age-induced marginal diagnosis. In particular, to analyze the role

[5] In all of our models, we control for the younger child's *own* relative age for grade because a large body of research documents that relative age for grade affects one's human capital, economic, and well-being outcomes (Bedard and Dhuey, 2006; McEwan and Shapiro, 2008; Elder and Lubotsky, 2009; Black et al., 2011; Kawaguchi, 2011; Fredriksson and Öckert, 2014; Hurwitz et al., 2015; Depew and Eren, 2016; Cook and Kang, 2016; Landersø et al., 2017; Dhuey et al., 2019).

[6] There could be other channels through which an older child's relative age for grade influences the outcomes of his/her household members—e.g., through changes in parental marital stability and maternal labor supply (Landersø et al., 2019) or sibling "role model" effects (Karbownik and Özek, 2019). However, these channels are less likely to explain the existence of ADHD spillovers across cousins *who do not share a household*. The most conservative interpretation of our results on educational and economic outcomes is that any possible benefits of marginal ADHD diagnosis spillovers do not come close to offsetting the negative impacts of being a younger relative of a youngest-for-grade child.

of schools, we use variation stemming from Swedish school financing rules that give schools different economic incentives to refer students for ADHD screenings,[7] and find that the sibling spillover effects are larger in municipalities with higher incentives. Moreover, the spillover effects among cousins provide some insights into the role of families. We find that ADHD spillovers exist even among cousin pairs that live in different municipalities and therefore do not attend the same schools or see the same health care providers, but have parents—who are siblings themselves—who likely communicate about their children's health issues. Thus, the ADHD spillovers that we document appear to in part materialize due to a "referral gap" between the younger family members of marginally-diagnosed and marginally-undiagnosed children.

Next, in the screening stage, the physician's diagnostic technology plays a critical role. For conditions for which there exists a precise technology that rules out erroneous or low-value diagnoses (e.g., a genetic test), "hereditary tagging" would only have the upside of improved targeting of screening resources, but no downside.[8] However, ADHD falls into a large class of health conditions for which there is no precise technology that can rule out erroneous or low-value diagnoses; instead, physicians have a noisy screening protocol. If the same (noisy) diagnostic criteria are applied to all younger relatives of previously diagnosed patients, the referral gap will translate into spillovers of marginal low-value—or potentially inaccurate—diagnoses. We provide empirical evidence suggesting that physicians indeed apply the same diagnostic criteria to all younger siblings of previously diagnosed patients, regardless of the older sibling's relative age for grade. This finding suggests that physicians adhere to the current protocol, which indicates that a prior diagnosis of ADHD in the family should be taken into account, *but which does not prescribe differentiating the significance of this information based on the previously diagnosed family member's relative age for grade.* In practice, however, when a physician is screening a younger relative of a previously diagnosed child, the older child's month of birth provides information about the expected severity of that child's condition and hence the strength of the signal of the hereditary component of the younger child's risk of ADHD. This points to a simple tool for mitigating the downsides of using family medical

[7] Similar incentives also exist in the U.S., see Cullen, 2003; Morrill, 2018.

[8] Specifically, the use of "hereditary tagging" would not result in any low-value or erroneous spillover diagnoses.

history in the ADHD diagnostic process: Adjusting the screening protocol so that it attaches less weight to information about a diagnosis of an older family member when he or she is relatively young-for-grade.

In sum, our two new empirical facts—(i) that marginal ADHD diagnoses propagate to younger family members, and (ii) that the value of these spillover diagnoses appears to be negligible (at least in terms of economic and human capital outcomes)—highlight an important downside of the use of family medical history in health care. Importantly, this does *not* imply that the use of family medical history should be eliminated—screening the family members of previously diagnosed individuals helps target scarce screening resources toward *ex ante* high-risk individuals. However, our analysis points to a simple modification that can equip the health care system to leverage the targeting benefits of "hereditary tagging," while reducing the potential costs of marginal diagnosis spillovers: Adjusting screening protocols so that a family member's prior diagnosis is given more weight when it signifies a higher expected risk of the condition in that individual.[9]

Our paper contributes to three strands of literature. First, the idea that marginal or over-diagnoses raise health care costs and sometimes adversely affect patients' well-being has been documented in a variety of settings (Brewer et al., 2007; Bond et al., 2013; Ong and Mandl, 2015; Mullainathan and Obermeyer, 2017; Einav et al., 2019; Chan et al., 2019; Alalouf et al., 2019; Obermeyer et al., 2019). Particularly relevant to our study of a mental health condition, Bos et al. (2020) show that marginal diagnoses of mental illness have adverse impacts on the future health and labor market outcomes of Swedish men in the military. We show that in settings where family history is used as a tag for further screening, these costs can be substantially *amplified*, as a single marginal diagnosis can spill over to multiple other family members. More broadly, our analysis uncovers an unintended consequence of using tags to target screening for a large set of medical conditions in which the diagnosing technology is noisy—e.g., this is an issue for a wide range of mental illnesses, see Frank and McGuire (2000); Anttila et al. (2018); Currie and Macleod (2020); Cuddy and Currie (2020)— the tag may propagate low-value (and potentially erroneous) diagnoses, and thereby the misallocation of

[9]The exact information needed to infer the expected risk of the condition in the family member depends on the condition. In the case of ADHD, it is sufficient to know the older child's month of birth. In the case of diabetes, the information could be gleaned from data on the relative's blood sugar level relative to the diagnosis threshold (Alalouf et al., 2019).

treatment, throughout society.

Second, our paper contributes to a burgeoning literature about the drivers of the increase in ADHD diagnoses in the last few decades (see, e.g., Chorniy et al., 2018). We document that one well-known process that generates marginal ADHD diagnoses—differences in maturity being interpreted as differences in ADHD—is amplified due to diagnosis spillovers throughout the family tree. The large magnitudes of our estimated spillover effects on younger siblings and cousins suggest that this mechanism can explain a sizeable share of the "exploding" caseloads of ADHD (Hinshaw and Scheffler, 2014).

Third, our results contribute to a growing body of evidence that establishes the family as an important nexus of the transmission of spillovers. Spillovers in non-health-related choices and outcomes are well documented.[10] A smaller literature analyzes how health-related interventions and health shocks to one child affect his/her siblings' cognitive skills and educational outcomes (see, e.g., Fletcher et al., 2012; Breining, 2014; Parman, 2015; Yi et al., 2015; Black et al., 2017; Breining et al., 2019), arguing that shifts in parental resource allocation across siblings may be an important mechanism. More closely related, Alsan (2017) finds that a Turkish national vaccination campaign targeting children under five years old has spillover effects on vaccine take-up among ineligible older siblings.[11] Our evidence on families' role in perpetuating ADHD diagnoses across siblings and cousins also relates to Chen et al. (2019), who document spillovers of medical information within the family tree. The key difference in our paper, however, is that the medical information that is transmitted across family members may be *de facto* incorrect.

2 Institutional Background

In this section, we describe the key institutional features of the Swedish health care and school systems that are relevant for our analysis of the spillovers of marginal diagnoses of ADHD.

[10] For evidence on sibling spillovers in test scores, educational attainment, college choice, and major choice, see, e.g., Aguirre and Matta (2018); Dustan (2018); Joensen and Nielsen (2018); Qureshi (2018a,b); Goodman et al. (2019); Karbownik and Özek (2019); Nicoletti and Rabe (2019); Altmejd et al. (2020); Dahl et al. (2020). For sibling spillovers in military service, see Bingley et al. (2019); for spillovers in program take-up (e.g., paternity leave), see Dahl et al. (2014).

[11] For other research on sibling spillovers in health outcomes, see also Altonji et al. (2017), who assess the extent to which the large sibling correlations in substance abuse are causal.

2.1 Screening Referral, Diagnosis, and Treatment of ADHD in Sweden

Sweden has a universal health care system in which the government operates as a large public insurer and finances its expenditures using tax revenue. Coverage includes inpatient care, primary and specialty outpatient care, and prescription pharmaceuticals.[12] Patients incur very low out-of-pocket costs, meaning that health care is effectively "affordable for all."[13]

The process of obtaining an ADHD diagnosis for a child involves several stages: Referral, screening, and diagnosis, which we describe below. Parents, schools, and physicians are all involved in this process.

Referral for ADHD screening. In order to be screened for ADHD, an individual needs to see a specialist provider (a psychiatrist). As discussed further below, referrals to screenings are often initiated by children's schools, although parents can also request referrals themselves.[14]

ADHD evaluation and diagnosis. An ADHD screening involves several components, described in detail in Socialstyrelsen (2014). First, using information from interviews with parents and teachers or other caregivers, the child is assessed using the Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association. An ADHD diagnosis requires six or more symptoms of hyperactivity and impulsivity or six or more symptoms of inattention, for children aged 16 or younger (from age 17, only five symptoms are required). Further, the symptoms need to be present in at least two settings, at home

[12] Some clinics and hospitals are run privately, but are incorporated into the public health care system and publicly funded. A subset of private clinics (also) serve patients who have a private health insurance policy on top of the universal public one. Healthcare is organized at the regional level, so there are some (usually minor) regional differences in coverage.

[13] An individual's maximum out-of-pocket spending for health care is approximately $120 per year. For prescription drugs, the maximum out-of-pocket spending per household is approximately ($247) over a rolling twelve-month window. For the purposes of calculating a household's total out-of-pocket drug spending, a household is defined as one adult plus all children aged 18 or below who reside in the same home.

[14] Demand for these screenings often exceeds supply, as evidenced by long waiting periods between the time of referral and the actual screening. While exact statistics on such waiting times are not available on an aggregate level, anecdotal evidence suggests that waiting times can be one year in duration. (`https://www.1177.se/behandling--hjalpmedel/undersokningar-och-provtagning/psykiatriska-utredningar/neuropsykiatrisk-utredning/`, accessed on November 11, 2020.)

and in school.[15] Second, the ADHD screening includes a physical exam and an evaluation of the child's family history of ADHD.[16]

ADHD treatment. Prescription drugs treating ADHD have been available in Sweden since 2002, when the first drug with the active substance Methylphenidate was permitted for treatment of ADHD in children below age 18.[17] Other active substances were subsequently authorized as well, and Sweden's National Board of Health and Welfare (NBHW) has documented a continuous and substantial increase in the rate of prescriptions of ADHD drugs since 2005 (Socialstyrelsen, 2012), which is the year when our prescription drug data begin. The NBHW also reports that both prevalence (share treated) and incidence (share initiating treatment) are highest among school-aged children (Socialstyrelsen, 2015). ADHD drugs can only be prescribed by psychiatrists in Sweden.

During our analysis sample period, ADHD drugs are only formally authorized for treatment of individuals below age 18 (in some cases, exceptions are made for individuals who initiate treatment before age 18 to continue treatment beyond that age). In practice, however, off-label prescriptions to individuals ages 18 and older are common in the data. A subset of the drugs authorized to treat ADHD are specifically recommended for children ages 5–18 in a national guideline issued by the NBHW.[18] In our analysis, we study both the broader set of ADHD drugs as well as the narrower set recommended for children.

Figure 1 plots the trend in ADHD diagnoses and drug treatment rates among children ages 6–19 over the years 2006 to 2017. Consistent with the rise in ADHD cases worldwide, the share of children who are diagnosed with ADHD has increased five-fold, while the share

[15] The DSM lists nine symptoms of hyperactivity and impulsivity, and nine symptoms of inattention. The DSM is revised continuously. See Appendix A for more information and the complete list of symptoms.

[16] Interestingly, Socialstyrelsen (2014) does not specify more precisely how a family history of ADHD should be incorporated into the diagnostic criteria, a fact that we return to in Section 5.4 below. Similarly, *UpToDate*, a service that aggregates medical research for clinical practice, states that "Family history of similar behaviors is important because ADHD has a strong genetic component," but does not provide more specific details on how the physician protocol for diagnosing ADHD should incorporate a family history of the condition. See: `https://www.uptodate.com/contents/attention-deficit-hyperactivity-disorder-in-children-and-adolescents-clinical-features-and-diagnosis?search=ADHD&topicRef=623&source=see_link`, accessed on November 9, 2020.

[17] Methylphenidate's trade names in the U.S. include Concerta, Methylin, Ritalin, and Equasym XL.

[18] As this guideline does not contain all molecules authorized for treatment of ADHD, it implicitly recognizes that there exist off-label prescriptions among adults. See Section 3 for exact Anatomical Therapeutic Chemical codes for all ADHD drugs, as well as for the subset of drugs that are recommended for children.

of children who are treated with ADHD drugs has increased six-fold over this time period in Sweden.

2.2 The Swedish School System

The school entry cutoff. During the period that we study, all children in Sweden start school in the fall of the year they turn seven years old; thus, the school entry cutoff is January 1.[19] The law requires children to attend school (*"skolplikt,"* which approximately translates to "school duty"), and homeschooling is not permitted. There exists some limited scope for "redshirting," or holding a child back by a year when starting school, though this practice is uncommon.[20]

School financing and ADHD detection. Swedish school financing rules generate economic incentives for schools to help detect and initiate the treatment of children's mental health conditions. Municipalities provide funding to schools—the law prohibits charging parents any direct school fees—and typically offer schools additional resources for students who have "special needs," which include mental health conditions like ADHD. Thus, school principals, teachers, and other administrators have an incentive to help parents detect and treat ADHD in their children. Indeed, Hjörne (2012) argues that most ADHD evaluations are initiated by teachers, who alert parents of their children's behavioral problems at school, and suggest that they seek further evaluation.

The Swedish school health care system (*Skolhälsovården*) facilitates this process. All children attending school in Sweden receive free annual health check-ups, and the most recent guidelines (issued in 2002) state that these check-ups must include evaluations of children's mental health and concentration skills in some years (Socialstyrelsen, 2002). The guidelines also state that all students have the right to additional evaluations for any health issues detected in these screenings at school. Thus, the school health care system plays an explicit role in referring children to specialists for a variety of issues, including ADHD.[21]

[19] The law has subsequently changed and children now start in the fall of the year they turn six.

[20] Because the school system is organized at the municipal level, exact statistics on the share of students who delay school entry are not available. That said, the law stipulates that a "delayed school duty" *(uppskjuten skolplikt)* is to be used sparingly, and must be authorized by the municipal school board after a formal application is made by the parents.

[21] In other countries, families may have their own incentives to obtain certain diagnoses for their children

3 Data and Sample

For our main analysis, we merge population register data from Statistics Sweden containing demographic and labor market information with outpatient and prescription drug claims data from the National Board of Health and Welfare. In some analyses, we also make use of birth records data from the NHBW. The population register data are available annually from 1990 to 2018, except information about the high school grade point average (GPA), which is available through 2016; outpatient records are available for years 2001–2016; prescription drug claims are available for the period July 2005 to December 2017; and birth records are available for years 1995–2016.

Two of our primary outcomes on ADHD treatment are measured using the prescription drug claims data. Therefore, we construct our main analysis sample such that our cohorts are old enough to be observed with prescriptions for ADHD medications. We begin with all individuals born in the years 1990–1996 who have at least one full younger sibling (born in 1990 or later). We drop families that have any non-singleton births. We only keep one sibling pair per family (i.e., for every individual born in 1990–1996, we only keep the next younger sibling). We also identify all of the first cousins of these sibling pairs, and then for each older sibling from the original sample, keep his/her next younger cousin.

The population register data have information on every individual's year and month of birth, which we use to construct the running variable in our RD analysis (described below in Section 4). While this information is available for all Swedish residents regardless of whether they are born in or outside of Sweden, Appendix Figure B1 makes it clear that individuals born outside of Sweden, Norway, and the European Union (EU) are disproportionately likely to have January recorded as their month of birth. Thus, to reduce the scope for measurement error in our running variable, we only keep sibling pairs in which both children are born in Sweden, Norway, or any of the 28 EU countries. We also only keep younger cousins who are born in these countries. Our final sample consists of 293,994 sibling pairs and 208,437 cousin pairs.

because they can then receive public benefits (see, e.g. Quadagno, 1997; Kubik, 1999 for evidence for the U.S.). In Sweden, however, a child's ADHD diagnosis does not confer any additional public assistance to families.

Key variables. We examine ADHD diagnoses using outpatient data, which includes visits to psychiatrists. Our main outcome is an indicator for having at least one outpatient claim with an ICD-10 code that starts with "F90" (the category for Attention Deficit Hyperactivity Disorders).

We study ADHD drug treatment using the prescription drug data. Our main outcome from the drug data is an indicator that is equal to one if a child has at least one claim for a drug used to treat ADHD ever observed in the prescription drug data. As an alternative outcome, we also consider an indicator for a claim based on a narrower definition of ADHD drugs, which are specifically recommended for children ages 5–18, as discussed in Section 2.[22] When coding the child-specific ADHD drug indicator, we further limit the observation window such that children born within the same fiscal year (July-June) are observed for the same number of months (see Appendix Table B1 for more details).

We additionally use the population register data to study educational and labor market outcomes of the younger siblings and cousins. We measure earnings in 2018, the most recent year of data that we can access.[23] We also create an indicator for ever having been enrolled in college by 2018, and analyze individuals' high school GPA (as reported in 2016, the last year that we observe this outcome). To ensure that our cohorts are old enough to measure these outcomes, we use a sample of sibling and cousin pairs in which the younger siblings and cousins are born in 1990–1995 only (i.e., they are aged 23–28 in 2018).

The population register data also provides us with a number of family-level control variables, including sibling/cousin birth spacing (in months), each child's gender, whether each parent is foreign-born, parental education level, and household income.[24]

Lastly, we use the birth records data for some supplementary analyses to test the identifying

[22]Specifically, for the broader ADHD drug definition, we consider all claims with Anatomical Therapeutic Chemical (ATC) codes that start with "N06BA" except "N06BA07", as well the ATC code "C02AC02". For the narrower ADHD drug definition, we use all claims with the following ATC codes: "N06BA04", "N06BA12", "N06BA09", "N06BA02", "C02AC02".

[23]We examine two measures of earnings: Employment earnings as well as a broader measure that additionally includes employment-based transfers such as parental leave benefits.

[24]While we also observe parental age and marital status, we do not include these control variables in our main specifications because these variables are recorded on an annual (calendar year) basis and thus exhibit a mechanical discontinuity between children born in December and January within any given fiscal year. For example, parents of children born in January are mechanically on average approximately one year older than parents of children born in December. However, as we show in Panel A of Appendix Table B10, our main results are unchanged when we add these controls.

assumptions of our RD model. These data contain information on exact dates of birth and gestation length for all children born in Sweden, allowing us to examine potential manipulation of the running variable in our RD models.

Sample means. Appendix Table B2 presents sample means of some of the key variables in our analysis. The first column uses our main sibling analysis sample, while the second and third columns are split into families with older siblings born in July-December and January-June, respectively. About 3.9 (4.4) and 4.8 (5.5) percent of older and younger siblings in our sample have an ADHD diagnosis (ever have an ADHD drug claim), respectively. About 4.6 (5.3) percent of the younger cousins have an ADHD diagnosis (ever have an ADHD drug claim). Mean birth spacing is about 36.6 months, approximately 15.3 percent of children have a foreign-born mother, while 11.6 percent have a mother with a college degree or higher. Families with children born in the two halves of the year are fairly similar in terms of observable characteristics, although rates of ADHD diagnosis and drug treatment are noticeably higher among children in families with July-December than January-June births. We explore these differences more formally in subsequent analyses.

4 Empirical Design

Our goal is to analyze how a marginal ADHD diagnosis among older children affects their younger siblings' and cousins' ADHD-related outcomes. To do so, we leverage the discontinuity in the older child's likelihood of own ADHD diagnosis and drug treatment generated by the difference in relative age for grade between children born just before and just after the Swedish school entry cutoff of January 1. As noted in Section 3, we do not have information on exact dates of birth for the main cohorts of older siblings in our analysis, and we therefore use the older child's month of birth relative to January as the running variable in our RD models.[25]

We begin by estimating an RD model to study the *own* relative age effect on ADHD diagnosis and drug treatment among the older siblings in our sample:

[25] Following Lee and Card (2008)'s guidance on RD estimation with a discrete running variable, we estimate a parametric RD model with standard errors clustered on the running variable (i.e., the older child's month of birth). Our results are also robust to using the wild cluster bootstrap to conduct inference (Cameron et al., 2008; Cameron and Miller, 2015), see Appendix Table B12.

$$ADHD_i = \alpha_0 + \alpha_1 \mathbf{1}[M_i \geq c] + f(M_i - c) + \mathbf{1}[M_i \geq c] \times f(M_i - c) + \mathbf{x'}_i\kappa + \theta_i^y + \epsilon_i \quad (1)$$

for every older sibling i. $ADHD_i$ is an ADHD-related outcome (i.e., an indicator for either a diagnosis or drug take-up). c denotes January, the school entry cutoff month. The variable $\mathbf{1}[M_i \geq c]$ is an indicator for the older sibling i being born in months January-June of every year, and zero otherwise. $f(M_i - c)$ is a linear function of the running variable, the older child's month of birth centered around January, which we allow to have different slopes on opposite sides of the cutoff. Since our running variable is centered around January, our cohorts represent fiscal (i.e., July-June) rather than calendar years. Thus, we control for fiscal year fixed effects, θ_i^y. The vector $\mathbf{x}_i$ includes the following controls: an indicator for whether the older child is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth (high school only, some college, college degree or more), and the log household income of the family averaged over the year of the older child's birth and the following two years. The main coefficient of interest is α_1, which represents the difference in ADHD diagnoses or drug treatment rates between older children who are born before and after January within every fiscal year.

Then, to study spillover effects on younger siblings' and cousins' outcomes, we estimate models of the following form:

$$\begin{aligned} Y_{ij} = \beta_0 &+ \beta_1 \mathbf{1}[M_i \geq c] + f(M_i - c) + \mathbf{1}[M_i \geq c] \times f(M_i - c) \\ &+ \beta_2 \mathbf{1}[M_j \geq c] + f(M_j - c) + \mathbf{1}[M_j \geq c] \times f(M_j - c) \\ &+ \mathbf{x'}_{ij}\pi + \rho_i^y + \lambda_j^y + \varepsilon_{ij} \end{aligned} \quad (2)$$

for each family with an older sibling i and a younger sibling or cousin j. Y_{ij} is an outcome of interest, such as an indicator for the younger child having an ADHD diagnosis. In addition to the variables capturing the older sibling's month of birth relative to January that are the same as in equation (1), we control for the younger sibling's/cousin's *own* relative age for grade by including analogous variables based on the younger child's month of birth centered around

January: $\mathbf{1}[M_j \geq c]$ and $f(M_j - c)$. We also control for fixed effects for every fiscal year of birth of the older and younger child, ρ_i^y and λ_j^y, respectively. The vector $\mathbf{x}_{ij}$ now includes the following family-level controls: sibling or cousin birth spacing (in months), indicators for whether the older and younger sibling/cousin is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth, and the log household income of the family averaged over the first three years of the older child's life. The main coefficient of interest is β_1, which represents the difference in outcomes between children in families where the older children are born before and after January in every fiscal year, holding constant the younger sibling's/cousin's birth month relative to January.

Identification and interpretation. The RD design relies on the assumption that only the treatment variable is changing discontinuously at the cutoff; all other variables possibly related to the outcomes we study should be continuous functions of the running variable (Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

In our regression models, the running variable is the older child's month of birth relative to January, and the treatment variable is an indicator for the older child being born in January–June and thus relatively older-for-grade than his/her counterparts born in the preceding six months of the same fiscal year. This treatment variable, in turn, generates a discontinuity in ADHD diagnoses among the older children. We discuss two potential issues for identification and interpretation in this setting: (1) Non-random sorting of families at the school-entry cutoff, and (2) non-ADHD-related channels through which an older child's relative age for grade might influence his/her younger family members' outcomes.

To assess issue (1), we begin by plotting a histogram of births at a daily level using birth records data in Appendix Figure B2. We use data on all children born in Sweden between January 1995 and December 2014 with information on exact date of birth, and consider a bandwidth of 180 days surrounding the January 1 cutoff. The figure makes it clear that there tend to be fewer births in December than in January, with noticeable dips during the December holiday season. The RD manipulation test (Cattaneo et al., 2018) yields a statistically significant $t-$statistic of 13.06.

To investigate this difference in the density of births between December and January fur-

ther, we use the same sample to check whether there are any systematic differences in average gestational age. If parents were trying to either delay or speed up childbirth through, for example, planned inductions or caesarean section deliveries, then there may be a discontinuity in observed gestational age at the cutoff. We plot the average length of gestation in days by the child's birth month for all children born between January 1995 and December 2014 with non-missing information on gestation length in Appendix Figure B3. There is no visible discontinuity in mean gestation length between December and January births. Column (1) of Appendix Table B3 reports results from an RD regression that uses gestation length in days as the outcome—the coefficient on the indicator for being born after the January cutoff is insignificant and very small in magnitude.

Importantly, the significant difference in the number of births between December and January only poses a concern if the sorting is systematically related to our outcomes of interest. To assess this possibility, we turn to our main analysis sample of siblings and study whether there are any differences in family background characteristics between December and January births. We do not observe any discontinuous changes in average birth spacing, the gender composition of older and younger siblings, or parental education levels in Appendix Figures B4, B5, and B6, respectively. Appendix Figure B7 plots *predicted* ADHD diagnosis and drug treatment indicators of the older and younger siblings, as well as the younger cousins, by the birth month of the older sibling. The predicted variables are constructed by regressing each of the ADHD outcomes on birth spacing (in months), indicators for whether the older and younger sibling/cousin is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth, and the log household income averaged over the year of the older child's birth and the following two years. If anything, it appears that predicted ADHD is slightly higher in families in which the older siblings are born in January than those in which the older siblings are born in December, which goes in the opposite direction of our main results. Columns (2)–(6) of Appendix Table B3 report results from RD regressions that use the family background characteristics as outcomes; we find no evidence of statistically significant discontinuities based on the older child being born before versus after the January cutoff.

In sum, our analysis of gestation length and family background characteristics does not

reveal any systematic discontinuities at the school-entry cutoff, and suggests that the sorting observed in Appendix Figure B2 is unlikely to bias our RD design. To further address the concern about sorting, we also conduct a robustness check in which we estimate a "doughnut-RD" model that omits all families with older children born in December or January, which yields similar results to those from our baseline models (see Appendix Table B11).

While it does not appear that there is any selection on observable characteristics, one may still be concerned about selection into different birth months across families with different underlying mental health issues. For instance, it could be that families with a higher propensity for ADHD are more likely to have all their children in December than in January, inducing a correlation between one sibling's relative age and the other sibling's likelihood of ADHD diagnosis. We investigate this possibility by checking whether the younger sibling's relative age for grade predicts the older sibling's ADHD-related outcomes. In Appendix Figure B8, we find no evidence of higher ADHD diagnoses or drug treatment among older siblings whose younger siblings are born in December versus January. We also estimate model (2), which includes both the older and younger sibling's relative age for grade variables, using the *older* child's ADHD outcomes as the dependent variables. Appendix Table B4 shows that while the older child's own relative age for grade predicts ADHD diagnoses and drug treatment (as we discuss further in Section 5.1 below), we find no consistent evidence of a discontinuity in the older sibling's ADHD outcomes based on the younger sibling's relative age for grade.[26]

When it comes to issue (2), we recognize that the spillover effects of an older sibling's relative age for grade on his/her younger sibling's outcomes could in principle operate through various family dynamics. For example, Landersø et al. (2019) document that mothers of children who are oldest for their grade are more likely to be employed when their children are 7 years old, and parents of oldest-for-grade children are more likely to remain married or cohabiting by the time their children are 15 years old. Karbownik and Özek (2019) propose that there is a "role model" effect—younger siblings of children who are oldest for their grade may be more likely to follow in their footsteps and experience better educational outcomes

[26]Out of the five regressions reported in Appendix Table B4, only one shows a significant spillover from younger to older siblings—the spillover effect on diagnoses among siblings who are three years apart in age or less. Since these siblings are close in age (and likely attend the same schools), it seems possible that the "hereditary tagging" mechanism could explain this spillover as in our main analysis. See Section 5.4 for more detailed discussion of the mechanisms underlying our spillover effects.

than their counterparts with older siblings who are youngest for their grade.[27]

However, these family dynamics are less likely to be relevant for relatives who do *not* share a household or the same set of parents. We therefore believe that our evidence of ADHD diagnosis spillovers across cousins assuages concerns about these other channels, although of course it is impossible to definitively rule out all other alternative explanations.

Moreover, because an older child's relative age for grade may influence his/her younger relatives' human capital and economic outcomes through other channels, we do not use the older child's relative age for grade as an instrument for the younger child's ADHD diagnosis in an instrumental variables model (as the exclusion restriction may not be satisfied). That said, the fact that we find large negative effects on these outcomes (see Section 5.3) implies that even if there were any economic benefits of the spillover ADHD diagnoses, they do not come close to offsetting the costs of being a younger relative of a youngest-for-grade child.

5 Results

We begin by using our main sample of older siblings to confirm prior evidence (e.g., Elder, 2010; Evans et al., 2010; Dalsgaard et al., 2012; Morrow et al., 2012; Zoëga et al., 2012; Halldner et al., 2014; Krabbe et al., 2014; Pottegård et al., 2014; Chen et al., 2016; Schwandt and Wuppermann, 2016; Layton et al., 2018; Whitely et al., 2018; Root et al., 2019; Furzer et al., 2020) that children who are youngest for their grade are more likely to be diagnosed with and treated for ADHD than those who are oldest for their grade. Next, we document spillover effects of the older children's relative-age-induced marginal diagnoses of ADHD on their younger siblings' and cousins' likelihoods of ADHD diagnosis and drug treatment. To assess whether there are any economic costs or benefits associated with these spillovers, we also study younger siblings' and cousins' educational and labor market outcomes. Lastly, we explore possible mechanisms and present several robustness tests.

[27] See also Goodman et al. (2019) and Altmejd et al. (2020) for evidence of sibling spillovers in college enrollment and college choice.

5.1 Own Relative Age Effects on Older Children's ADHD Diagnoses and Drug Treatment

Figure 2 uses our main siblings sample and plots ADHD-related outcomes of the older siblings by their *own* birth month (centered around January). Sub-figure (a) plots the share of children with an ADHD diagnosis in the outpatient data, while sub-figure (b) plots the share of children with at least one ADHD drug claim in the prescription drug data. Both graphs show clear discontinuities between December and January births—children who are youngest for their grade (i.e., born in December) are substantially more likely to be diagnosed with ADHD and to use ADHD prescription drugs than those who are oldest for their grade.

Table 1 reports results from estimating model (1) and confirms the graphical evidence. The outcome in columns (1) and (2) is an indicator equal to one if the older sibling ever has an outpatient claim with an ADHD diagnosis. The outcome in columns (3) and (4) is an indicator equal to one if the older sibling has at least one ADHD drug claim. In columns (2) and (4), the main sample is further limited to sibling pairs with an age difference of no more than 36 months. The outcome in column (5) is an indicator equal to one if the older sibling has at least one ADHD drug claim that is specifically recommended for children between ages 5 and 18. In addition, the observation period for drug claims in column (5) is restricted such that children born in the same fiscal year are observed for the same number of months (see Appendix Table B1).

Across all columns, we find that being born after the school entry cutoff is associated with a significantly lower likelihood of being diagnosed with and treated for ADHD. Specifically, focusing on the estimates in columns (1) and (3), we observe that children born after the cutoff are 1.2 and 1.3 percentage points less likely to be diagnosed with ADHD and treated with ADHD drugs, respectively. Relative to the corresponding sample means, these estimates yield large effect size magnitudes of 30.8 and 29.5 percent, respectively.[28]

[28]Note that here, and in all of our regression tables, we report the coefficients on indicators for being born *after* the school entry cutoff (i.e., the effects of being the oldest-for-grade child). In our introduction, we instead discuss our results in terms of the impacts of being the youngest-for-grade child (or the sibling/cousin of a youngest-for-grade child), i.e., flipping the signs on these coefficients.

Interpreting the ADHD gap. To interpret this gap, it is helpful to consider the etiology of ADHD. Unlike conditions for which there is a precise screening mechanism that yields a discrete outcome—e.g., an X-ray can determine whether or not someone has a broken bone; a genetic test can identify women who are BRCA-gene positive or not—ADHD, like many mental health conditions, is diagnosed differently.[29] As noted by Levy et al. (1997), "ADHD is best viewed as the extreme of a behavior that varies genetically throughout the entire population rather than a disorder with discrete determinants." Put differently, ADHD symptoms—such as immaturity, impulsiveness, and attentiveness—vary naturally within the population, and an ADHD diagnosis is given to individuals whose symptoms fall in the tails of these distributions.

With this understanding of ADHD, what can we interpret about the additional diagnoses among December-born children relative to January-born children? Figure 3 presents a stylized visual framework to aid the interpretation of the ADHD diagnosis gap at the school-entry cutoff. The bell curves represent the distributions of underlying ADHD symptoms in the populations of children born in December and January, respectively. The yellow areas under each of the curves signify the children who receive positive ADHD diagnoses. This framework highlights that among both groups, children with the most extreme ADHD traits are diagnosed. However, on the margin, *children with lighter symptoms are diagnosed if they are born in December, but not if they are born in January.* This observation has an immediate implication, to which we return in our discussion of how to mitigate spillovers from "hereditary tagging" in Section 5.4: If we restrict attention only to children who are diagnosed with ADHD, then the average severity of ADHD will be higher among January-born than December-born children. This follows directly from the fact that the "compliers"—children who are diagnosed because they are born in December but would not have been diagnosed if they were born in January—have the lightest ADHD symptoms; that is, they are "on the margin."

The extensive literature that documents the relative-age-induced diagnosis gap in various contexts is often cited as evidence that ADHD is a commonly *over*-diagnosed condition, and that over-diagnoses contribute to the substantial rise in ADHD caseloads over the last few decades (Bruchmüller et al., 2012; Hinshaw and Scheffler, 2014; Schwarz, 2017; Merten et

[29] This is a fact that we discuss in more detail in Section 5.4 below, as the diagnosis technology at the physician's disposal has implications for whether "hereditary tagging" can entail adverse consequences, and whether such consequences can be mitigated.

al., 2017). In particular, a popular interpretation of the diagnosis gap is that immaturity is mis-classified as ADHD among children who are the youngest in the classroom and have more difficulties paying attention and sitting still than classmates who are nearly one year older. Thus, youngest-for-grade children are over-diagnosed with ADHD. By the same logic, however, children who are oldest in the classroom may be under-diagnosed due to their relative maturity when compared to their (younger) peers in the classroom.[30]

More generally, the simple framework described in Figure 3 highlights that there are in fact several possible interpretations of the ADHD diagnosis gap: (i) Over-diagnosis among December-born children, potentially coupled with under-diagnosis among January-born (the most common interpretation, as just discussed); (ii) over-diagnosis among all children, but more among December-born; and (iii) under-diagnosis among all children, but less among December-born. As illustrated in Appendix Figure B9, the "correct" interpretation depends on the "true" ADHD cutoff, i.e., the threshold in the distribution of ADHD symptoms at which the scientific community decides to define an individual as having ADHD. This is something that is unobserved in our data (or any other data set available to researchers), and we therefore do not take a stance on the direction of error represented by the diagnosis gap.

Instead, what is key to our empirical design is that the school entry cutoff generates additional marginal diagnoses among December-born children, and, under all three interpretations, December-born children with lighter symptoms are diagnosed on the margin. Thus, in the rest of the analysis, we use the discontinuity in the diagnosis rate between children born on opposite sides of the school-entry cutoff to study spillover effects of marginal diagnoses on younger family members.[31]

[30] Age-for-grade is not the only characteristic with respect to which there may be over- or under-diagnosis. For example, some studies point to the risk of under-diagnosis of ADHD, especially among girls (Visser, 2014; Furzer et al., 2020), and demonstrate heterogeneity in the types of diagnostic errors with respect to child gender, race, and socioeconomic status (Furzer, 2020; Marquardt, 2020).

[31] One other alternative interpretation is that children who are young-for-grade develop ADHD as a consequence of being youngest in their grade (i.e., the development of ADHD is endogenous to one's relative age for grade). Such a scenario would imply that the distribution of ADHD risk in our figures should be shifted to the right among December-born children. If so, the diagnosis gap would not reflect differential rates of diagnoses for the same severity of ADHD symptoms, but instead a higher share of children with sufficiently severe ADHD traits among children who are young-for-grade. Thus, conditional on diagnosis, we would not observe that children who are young-for-grade have lighter symptoms, on average. This contrasts with evidence from, e.g., Furzer (2020), who finds that youngest-for-grade children in Canada have a relatively lower risk of mental illness. Further, while the diagnosis gap is present in many contexts, there exist settings in which it does not (e.g., Dalsgaard et al., 2012), which would be inconsistent with younger-for-grade children developing ADHD symptoms.

5.2 Spillover Effects on ADHD Diagnoses and Drug Treatment of Younger Siblings and Cousins

In Figure 4, we present graphical evidence that a younger sibling's likelihood of ADHD diagnosis and drug treatment depends on his/her older sibling's relative age for grade. The figure is analogous to Figure 2, except that it plots the younger sibling's ADHD diagnosis and drug treatment rates on the respective sub-figure $y-$axes. We find that younger siblings of older children born in December are more likely to be diagnosed with ADHD and to be treated with ADHD drugs than their counterparts with older siblings born in January.

Table 2 presents the corresponding regression estimates of model (2) for the same outcomes as in Table 1, except that they are now measured for the younger sibling. Not surprisingly, the younger sibling's *own* relative age for grade has an effect on the probability of ADHD diagnosis and drug treatment—columns (1) and (3) show that younger siblings born after the school entry cutoff are 1.6 percentage points less likely to have an ADHD diagnosis or an ADHD drug claim than those born before the cutoff. However, conditional on the younger sibling's own relative age for grade, we also see a significant effect of the older sibling's relative age for grade on the younger sibling's likelihood of ADHD diagnosis and drug treatment. Younger siblings of older children who are born after the school entry cutoff are 0.6 percentage points less likely to have an ADHD diagnosis or an ADHD drug claim, corresponding to 12.0 and 9.8 percent effect sizes when evaluated at the respective dependent variable means. The magnitudes of the spillovers on ADHD diagnoses and drugs are 38.1 and 33.5 percent, respectively, of the sizes of the younger child's own relative age-for-grade effects on these outcomes.[32]

As noted previously, an older child's relative age for grade could in principle influence a younger sibling's outcomes through a variety of family dynamics. However, such alternative mechanisms are unlikely to be relevant for family members who do not share the same household. To that end, we examine spillover effects on ADHD-related outcomes among first cousins. Figure 5 plots ADHD-related outcomes of the younger cousins by their older cousin's

[32] Appendix Table B5 examines heterogeneity in the spillover effects across sibling pairs with differing gender composition. We find that the spillovers on ADHD diagnoses are somewhat larger in sibling pairs with at least one girl, while the spillovers on ADHD drugs are stronger in sibling pairs with at least one boy. However, the coefficients are not all significantly different from one another, suggesting that these spillovers exist regardless of the siblings' gender mix. We also explored heterogeneity with respect to various measures of the families' socioeconomic status, finding no consistent patterns.

birth month (i.e., the birth months of the older siblings from our main sample). Sub-figure (a) shows the graph for ADHD diagnoses in the outpatient data, while sub-figure (b) plots ADHD drug treatment rates. We find that younger cousins whose older cousins are born in December are more likely to be diagnosed with ADHD and to be treated with ADHD drugs than those whose older cousins are born in January. Table 3 reports the corresponding regression estimates—we find that younger cousins whose older cousins are born after the school entry cutoff are 0.6 percentage points (13.0 percent) and 0.5 percentage points (9.1 percent) less likely to be diagnosed with ADHD and to be treated with ADHD drugs, respectively, than their counterparts whose older cousins are born before the cutoff. These spillovers amount to 32.1 and 24.2 percent of the younger cousins' own relative age-for-grade effects on ADHD diagnoses and drug treatment, respectively.

5.3 The Economic Impacts of Spillovers of Marginal ADHD Diagnoses

What are the economic implications of the spillovers of marginal ADHD diagnoses? To help shed light on this question, we examine spillover effects on younger siblings' and cousins' economic and educational outcomes. In Figures 6 and 7, and Table 4, we analyze the younger siblings' and cousins' earnings (using two measures described in Section 3), high school GPA, and college enrollment. As noted in Section 3, we use a sample of siblings and cousins who are born in 1990–1995 in these analyses, and measure the labor market outcomes in the latest year of data that we have. Thus, our cohorts are 23 to 28 years old at the time of observation. We also observe whether they are ever enrolled in college through 2018, and measure their high school GPA in 2016.

Consistent with the existing evidence of the effect of relative age for grade on children's *own* human capital attainment (Bedard and Dhuey, 2006; McEwan and Shapiro, 2008; Elder and Lubotsky, 2009; Black et al., 2011; Kawaguchi, 2011; Fredriksson and Öckert, 2014; Hurwitz et al., 2015; Depew and Eren, 2016; Cook and Kang, 2016; Landersø et al., 2017; Dhuey et al., 2019), Table 4 shows that younger siblings and cousins who are born after the cutoff have a higher high school GPA than those born before the cutoff.[33] We also see negative effects

[33]At the same time, we see a negative effect of own relative age on college enrollment for younger siblings

of own relative age for grade on both of the measures of labor market earnings. These labor market effects of *own* relative age for grade are largely mechanical, however, because these outcomes are measured on an annual basis—within each fiscal year, relative to children born before the school entry cutoff, children who are born after the cutoff graduate one year later and therefore have less work experience in any given (calendar) year of observation.

When we examine spillovers on younger siblings and cousins conditional on own relative age for grade, such mechanical effects are no longer pertinent.[34] Table 4 shows that younger siblings whose older siblings are born after the cutoff have 1.6 percent higher labor market earnings and are 2.1 percent more likely to have ever enrolled in college than those whose older siblings are born before the cutoff (Columns 1 and 5, Panel A of Table 4). We also find that younger cousins of children who are born after the cutoff have a 1.0 percent higher GPA and are 2.6 percent more likely to have ever enrolled in college than the younger cousins of children born before the cutoff.

Our analysis of educational and labor market outcomes suggests that the younger siblings and cousins of December-born children tend to be worse off than the younger family members of January-born ones. While, as noted previously, we cannot definitively rule out all other channels by which an older child's relative age for grade influences his/her younger relatives' outcomes, these results suggest that there are no clear economic benefits associated with marginal ADHD diagnosis spillovers. If anything, the most conservative interpretation of our results is that even if there were some benefits of marginal ADHD diagnosis spillovers, they do not come close to offsetting the costs of being the younger relative of a youngest-for-grade child.[35]

Moreover, our results on educational and economic outcomes are complementary to the existing clinical evidence on the impacts of ADHD treatment. A positive ADHD diagnosis

and an insignificant relationship between own relative age and college enrollment for younger cousins.

[34]Moreover, all of our models account for the effects of age on the educational and labor market outcomes measured in 2018 (or 2016 in the case of high school GPA) because we control for the running variable for both the older and younger child (i.e., month of birth), as well as fiscal year of birth fixed effects for each child. Thus, our estimated impacts are not driven by the mechanical difference in age between January- and December-born children when they are observed in our outcome data.

[35]It is also possible that having an older child in the family who is more likely to be treated with ADHD medication makes the home environment calmer for everyone else. Thus, we might expect that younger siblings of children who are youngest for their grade actually fare better than their counterparts whose older siblings are oldest for their grade. This type of mechanism would imply that our spillover effects are biased downward.

typically leads to long-term use of ADHD prescription drugs—50 percent of patients who initiate ADHD drugs remain on them five years later (Socialstyrelsen, 2012). Existing evidence on the impacts of using ADHD drugs is mixed and limited to studies of selected short- and medium-term behavioral and educational outcomes.[36] To date, we know very little about the long-run impacts of ADHD drug treatment on measures of individuals' well-being, especially when it comes to economic outcomes.[37] Our estimates provide suggestive evidence that there could be important long-term costs associated with the use of ADHD drugs on the margin.

5.4 Understanding the Mechanisms Behind Marginal ADHD Spillovers

We have established that there are substantial spillovers of marginal ADHD diagnoses to younger siblings and cousins, and that there appear to be no economic benefits (and may even be costs) associated with these spillovers. Now, we turn to an investigation of the underlying mechanisms generating these spillovers. This analysis, in turn, informs a discussion about whether such spillovers can be mitigated.

As described in Section 2, the ADHD diagnosis process involves two key steps: First, one must get a referral for an ADHD evaluation, and second, a physician performs a screening (which may result in a diagnosis). The existence of spillovers means that somewhere in this two-step process, a marginal diagnosis of an older child affects the likelihood that a younger family member is diagnosed. Next, we seek to understand where and how this happens.

5.4.1 Medical Family History in the Referral Stage

If information about an individual's medical family history is used in the referral stage, then a child with a family history of ADHD should have a higher likelihood of being referred to an ADHD screening than a child without such a family history. Indeed, the referral stage is one in which "hereditary tagging" can be used to allocate scarce screening resources toward *ex-ante* high-risk individuals.

[36] See, e.g., Jensen, 1999; Wilens et al., 2003; Charach et al., 2004; Dalsgaard et al., 2012; Humphreys et al., 2013; Molina et al., 2013; Currie et al., 2014; Chorniy and Kitashima, 2016; Cortese et al., 2018. Moreover, a child's positive mental health diagnosis may impose stigma costs and result in unfavorable expectations from teachers and school administrators (Moses, 2010; Ohan et al., 2011; Bharadwaj et al., 2017).

[37] A search of *UpToDate*®, a website that synthesizes medical research and is widely used by clinicians to inform treatment decisions, yields no information about the long-term effects of ADHD drug treatment on adult economic outcomes, such as earnings.

In our setting, "hereditary tagging" in the referral stage would imply that a relative-age-induced marginal diagnosis of one child will influence the likelihood that a younger sibling or cousin gets referred to an ADHD screening. As both schools and parents play important roles in requesting ADHD screenings, we investigate the roles of each of these two parties.

The role of schools. As described in Section 2, teachers and school administrators can refer children to psychiatrists for ADHD evaluations; further, schools often have direct economic incentives to detect ADHD in their students. Since siblings often attend the same school, teachers and administrators may be more likely to refer the younger siblings of older children who have ADHD diagnoses than the younger siblings of older children who do not. This practice may thus propagate marginal diagnoses across siblings.

While we do not observe schools' referrals to psychiatrists in our data, we conduct two empirical tests to shed light on the potential role of schools in generating marginal ADHD spillovers. First, in Appendix Table B6, we analyze whether the spillover effects on younger siblings' ADHD outcomes are larger or smaller when they are observed before or after the child graduates high school. Intuitively, if schools play an important role in perpetuating marginal diagnoses, then we would expect larger spillovers onto younger siblings while those younger siblings are still in school. Indeed, we find that the spillover effects on younger siblings' ADHD diagnoses and drug claims are stronger when they happen before the child graduates high school. However, the spillover on ADHD drug treatment also exists beyond high school, suggesting that schools may not be the only places where an older child's ADHD diagnosis leads to a higher likelihood of referral for the younger sibling.

Second, we explore whether the spillover effect is greater in municipalities in which schools have larger financial incentives to identify ADHD in their students. In Appendix Table B7, we examine heterogeneity in the sibling spillover effect across municipalities with a higher versus lower share of school funding targeted for special needs education. We augment model (2) with municipality fixed effects as well as with interaction terms between the RD variables and the share of funding for special needs as measured in the mother's municipality of residence when her youngest child is aged seven years old. The negative coefficients on the interaction terms indicate that the sibling spillover effects are larger when special needs funding is higher,

which is consistent with schools responding to funding incentives by identifying more children for possible ADHD diagnoses.[38]

In sum, both of these empirical tests suggest that schools may play a role in generating spillovers in the referral stage.

The role of families. In addition, families may also perpetuate marginal diagnoses in the referral stage. This would occur if an older child's (marginal) diagnosis raises the likelihood that the family requests a screening for the younger child as well.[39]

The existence of a spillover effect among first cousins provides some indication that families play a role. Cousins are less likely to attend the same school than siblings within one nuclear family, but their parents could communicate about their children's ADHD diagnoses and treatments.

To further assess the role of families relative to the role of schools, Appendix Table B8 presents separate estimates of the spillover effects across cousin pairs that reside in the same versus different municipalities. It appears that the spillover effect is somewhat stronger for cousin pairs that reside in the same municipality (and thus could potentially attend the same schools) than those who live in different municipalities. At the same time, for cousin pairs who are no more than 36 months apart in age, the spillover effects are similar regardless of whether they live in the same or different municipalities. Taken as a whole, these results suggest that both schools and families likely play a role in perpetuating the spillover effects of marginal ADHD diagnoses.

5.4.2 Medical Family History in the Diagnosis Stage

The above analysis suggests that schools and families are more likely to refer to screenings the younger siblings and cousins of older children who are marginally diagnosed with ADHD than the younger family members of older children who are not. Thus, as a consequence of such "hereditary tagging" at the referral stage, more children with an older relative born in

[38] See Cullen (2003) and Morrill (2018) for evidence on schools responding to financial incentives to classify students as having ADHD in the United States.

[39] There are multiple reasons for why an older child's diagnosis reduces the costs of the family requesting an ADHD screening for the younger child. For example, families in which one child is being treated for ADHD have an established connection with a child psychiatrist. In addition, drug treatment of a one child may make the parents less worried about side effects and thus less hesitant to seek treatment for another child.

December end up in the doctor's office for an ADHD screening than children with an older relative born in January.

At this point, understanding the physician's diagnostic technology is important. If there existed a technology that could precisely identify children as being ADHD-positive or negative (e.g., as is the case in BRCA genetic screening), then regardless of the fact that the younger family members of December-born children may be over-referred to ADHD screenings compared to the younger family members of January-born children, the physician would simply use the technology to accurately diagnose all children who show up in her office. Thus, "hereditary tagging" would only have an upside—improved targeting of scarce screening resources—but no downside in terms of spillovers of low-value or even erroneous diagnoses.

However, ADHD falls into a large class of health conditions for which there is no discrete test or diagnostic procedure that allows the physician to precisely determine which patients do and do not have the condition. Instead, physicians have a noisy screening protocol. Then, if the same (noisy) diagnostic criteria are applied to all younger relatives of previously diagnosed patients, the referral gap will translate into spillovers of marginal low-value—or potentially inaccurate—diagnoses. Thus, when the diagnosis technology is noisy, the practice of "hereditary tagging" can generate costs.

In Appendix Table B9, we provide empirical evidence that physicians apply the same diagnostic criteria to all younger siblings of previously diagnosed patients, regardless of the older child's relative age for grade. We estimate our main regression model (2), using a sub-sample of sibling pairs in which the older child is diagnosed with ADHD. In this sub-sample, we find no discontinuity in the younger child's likelihood of ADHD diagnosis based on whether the older child is oldest- or youngest-for-grade. In other words, conditional on the older child's diagnosis, physicians are equally likely to diagnose their younger siblings, regardless of the older sibling's relative age for grade. Importantly, this finding confirms that physicians correctly follow the current protocol. As discussed in Section 2, while the protocol prescribes that the presence of ADHD in the family is taken into account in the diagnostic process, it does *not* indicate that the physician should attach differential weight to this information depending on the previously diagnosed family member's relative age for grade.

However, recall that Figure 3 highlights that among children who are diagnosed with

ADHD, the average severity of ADHD symptoms should be lower among December-born than January-born children. This follows from the fact that the "compliers"—the children who are diagnosed because they are December-born but who would not have been diagnosed if they were January-born—have the weakest ADHD symptoms; that is, they are "on the margin" of diagnosis.

Thus, when a physician is screening a younger relative of a previously diagnosed child, the older child's month of birth provides information about the expected severity of that child's condition and hence the strength of the signal of the hereditary component of the younger child's risk of ADHD. A positive ADHD diagnosis of an older child who is born in December implies a *lower* expected ADHD risk of the younger child relative to a positive diagnosis of an older child who is born in January.

This discussion suggests that there exists a simple adjustment to physician protocol that can mitigate marginal ADHD spillovers. We propose that physicians should take into account information about the older child's relative age for grade, and attach less weight to the older child's ADHD diagnosis if that child is youngest-for-grade than if that child is oldest-for-grade. More broadly, our findings suggest that designing physician protocols that take into account whether a previously diagnosed relative's diagnosis may be marginal—and thus the strength of the signal of *ex ante* hereditary risk that stems from family medical history—could reduce the costs of marginal diagnosis spillovers, while allowing the health care system to leverage the targeting benefits of "hereditary tagging."

5.5 Additional Results

We perform a series of additional analyses to examine the sensitivity of our estimates. First, as noted in footnote 24, our baseline models do not include controls for parental characteristics that are measured on an annual level—age and marital status—because of the mechanical discontinuity they exhibit when comparing December- and January-born children. However, Panel A of Appendix Table B10 demonstrates that our results on spillover effects on younger siblings' and cousins' ADHD outcomes are robust to including them. In Panel B of Appendix Table B10, we also document that our results are largely unchanged if we use a quadratic polynomial in the running variable when estimating the RD models.

In addition, to further address concerns about potential sorting of families across birth months, we estimate "doughnut-RD" specifications that drop all families with older children born in December or January of any year. As we show in Appendix Table B11, our results remain similar with this sample limitation.

Finally, we examine the sensitivity of our results to an alternative method of inference. Appendix Table B12 demonstrates that our estimated spillover discontinuities remain significant when using the wild cluster bootstrap (Cameron et al., 2008; Cameron and Miller, 2015).

6 Conclusion

Growing evidence suggests that patients who are diagnosed "on the margin"—i.e., they would not have been diagnosed if there were a small change in their underlying symptoms or indicators of the disease—do not appear to be better off as a result of the diagnosis. In some cases, these patients may even be worse off than if they had not been diagnosed. Since diagnosed patients usually receive medical treatment, this means that marginal diagnoses increase health care spending without clear benefits for patients. Thus, understanding the drivers of low-value marginal diagnoses and mitigating their spread is an important goal for health policy.

At the same time, a large class of conditions have a hereditary component in their etiology, and information about family members' prior diagnoses is used to "tag" patients for screening as well as in the diagnostic process. While such "hereditary tagging" is more efficient than screening individuals at random, our paper uncovers an important cost of this common health care practice—the propagation of marginal, low-value diagnoses across family members. We focus on the case of ADHD, which is the most common mental disorder among children, and for which there exists a well-known determinant of marginal diagnoses—children's relative age for grade. We use Swedish administrative data and an RD design to show that children who are born shortly before the school entry cutoff and are youngest for their grade are 30.8 and 29.5 percent more likely to be diagnosed with ADHD and to be treated with ADHD drugs, respectively, than their oldest-for-grade peers born after the cutoff.

We then study the spillover effects of these marginal ADHD diagnoses on the focal children's younger siblings and cousins. We find that younger siblings of children born before

the cutoff are 12.0 and 9.8 percent more likely to be diagnosed with and treated for ADHD, respectively, than the younger siblings of children born after the cutoff. For younger cousins, the corresponding spillover effect magnitudes are 13.0 and 9.1 percent, respectively. These effects exist conditional on the younger children's own relative age for grade.

To investigate the long-term implications of these diagnosis spillovers, we also show that younger siblings and cousins of children born before the school entry cutoff have worse economic and human capital outcomes than the younger siblings and cousins of children born after the cutoff. While we cannot completely rule out that other changes in family behaviors associated with the older child's relative age for grade contribute to these effects, our results suggest that there are no clear benefits and may even be important costs of ADHD diagnoses induced by the marginal diagnoses of older family members.

Additional analyses suggest that schools and families play important roles in the "tagging process" at the referral stage, in which younger relatives of previously diagnosed children are systematically more likely to be referred for ADHD screenings. Moreover, although physicians follow protocol by incorporating information about the older child's ADHD diagnosis in their diagnostic criteria, they do not undo the "referral gap" because they do not take into account the older child's relative age for grade. Our analysis suggests that low-value marginal diagnosis spillovers could be mitigated with a small adjustment to this protocol that would assign less weight to information about the older child's diagnosis if that child is young-for-grade rather than old-for-grade.

Our evidence of large family spillover effects of marginal ADHD diagnoses also helps explain the rapid increase in ADHD caseloads both in the United States and in other countries. Our results underscore that a single marginal diagnosis can trigger the diagnoses of multiple other family members, thus spreading them rapidly throughout the population. Further research is needed to understand how these processes affect the propagation of diagnoses of many other medical conditions that have noisy diagnosing technologies and in which links between individuals are used for targeting screening.

References

Aguirre, Josefa and Juan Matta, "Walking in your footsteps: Sibling spillovers in higher education choices," 2018. Available at SSRN 3255875.

Alalouf, Mattan, Sarah Miller, and Laura R Wherry, "What Difference Does a Diagnosis Make? Evidence from Marginal Patients," Working Paper 26363, National Bureau of Economic Research October 2019.

Alsan, Marcella, "The Gendered Spillover Effect of Young Children's Health on Human Capital: Evidence from Turkey," Working Paper 23702, NBER 2017.

Altmejd, Adam, Andres Barrios-Fernandez, Marin Drlje, and Christopher Neilson, "Siblings' Spillover Effects on College and Major Choice: Evidence from Chile, Croatia and Sweden," 2020. Princeton University, unpublished manuscript.

Altonji, Joseph G., Sarah Cattan, and Iain Ware, "Identifying Sibling Influence on Teenage Substance Use," *Journal of Human Resources*, 2017, *52* (1), 1–47.

Anttila, Verneri, Brendan Bulik-Sullivan, Hilary K Finucane, Raymond K Walters, Jose Bras, Laramie Duncan, Valentina Escott-Price, Guido J Falcone, Padhraig Gormley, Rainer Malik et al., "Analysis of shared heritability in common disorders of the brain," *Science*, 2018, *360* (6395), eaap8757.

Barkley, Russell A, "Attention-Deficit/Hyperactivity Disorder," in D.A. Wolfe and E.J. Mash, eds., *Behavioral and emotional disorders in adolescents: Nature, assessment, and treatment*, Guilford Publications, 2006.

Bedard, Kelly and Elizabeth Dhuey, "The persistence of early childhood maturity: International evidence of long-run age effects," *The Quarterly Journal of Economics*, 2006, *121* (4), 1437–1472.

Bharadwaj, Prashant, Mallesh M Pai, and Agne Suziedelyte, "Mental health stigma," *Economics Letters*, 2017, *159*, 57–60.

Bingley, Paul, Petter Lundborg, and Stéphanie Vincent Lyk-Jensen, "Brothers in Arms: spillovers from a draft lottery," *Journal of Human Resources*, 2019, pp. 0317–8646R3.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "Too young to leave the nest? The effects of school starting age," *The Review of Economics and Statistics*, 2011, *93* (2), 455–467.

_ , Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen, "Sibling spillovers," Working Paper w23062, National Bureau of Economic Research 2017.

Bond, Mary, Toby Pavey, Karen Welch, Chris Cooper, Ruth Garside, Sarah Dean, and Christopher J Hyde, "Psychological consequences of false-positive screening mammograms in the UK," *BMJ evidence-based medicine*, 2013, *18* (2), 54–61.

Bos, Marieke, Andrew Hertzberg, and Andres Liberman, "Are we Overdiagnosing Mental Illnessses? Evidence from Randomly Assigned Doctors," 2020. Federal Reserve Bank of Philadelphia, unpublished manuscript.

Breining, Sanni, Meltem Daysal, Marianne Simonsen, Sandra E Trandafir Mircea, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen, "Spillover Effects of Early-Life Medical Interventions," Working Paper 9086, IZA Working Paper 2019.

Breining, Sanni Nørgaard, "The presence of ADHD: Spillovers between siblings," *Economics Letters*, 2014, *124* (3), 469–473.

Brewer, Noel T, Talya Salz, and Sarah E Lillie, "Systematic review: the long-term effects of false-positive mammograms," *Annals of internal medicine*, 2007, *146* (7), 502–510.

Bruchmüller, Katrin, Jürgen Margraf, and Silvia Schneider, "Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis.," *Journal of consulting and clinical psychology*, 2012, *80* (1), 128.

Cameron, A Colin and Douglas L Miller, "A practitionerâs guide to cluster-robust inference," *Journal of human resources*, 2015, *50* (2), 317–372.

_ , Jonah B Gelbach, and Douglas L Miller, "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

Chan, David C, Matthew Gentzkow, and Chuan Yu, "Selection with variation in diagnostic skill: Evidence from radiologists," Working Paper w26467, National Bureau of Economic Research 2019.

Charach, Alice, Abel Ickowicz, and Russell Schachar, "Stimulant treatment over five years: adherence, effectiveness, and adverse effects," *Journal of the American Academy of Child & Adolescent Psychiatry*, 2004, *43* (5), 559–567.

Chen, Mu-Hong, Wen-Hsuan Lan, Ya-Mei Bai, Kai-Lin Huang, Tung-Ping Su, Shih-Jen Tsai, Cheng-Ta Li, Wei-Chen Lin, Wen-Han Chang, Tai-Long Pan et al., "Influence of relative age on diagnosis and treatment of attention-deficit hyperactivity disorder in Taiwanese children," *The Journal of pediatrics*, 2016, *172*, 162–167.

Chen, Yiqun, Petra Persson, and Maria Polyakova, "The Roots of Health Inequality and the Value of Intra-Family Expertise," Working Paper 25618, NBER 2019.

Chorniy, Anna and Leah Kitashima, "Sex, drugs, and ADHD: The effects of ADHD pharmacological treatment on teens' risky behaviors," *Labour Economics*, 2016, *43*, 87–105.

_ , Janet Currie, and Lyudmyla Sonchak, "Exploding asthma and ADHD caseloads: The role of medicaid managed care," *Journal of Health Economics*, 2018, *60*, 1–15.

Cook, Philip J and Songman Kang, "Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation," *American Economic Journal: Applied Economics*, 2016, *8* (1), 33–57.

Cortese, Samuele, Nicoletta Adamo, Cinzia Del Giovane, Christina Mohr-Jensen, Adrian J Hayes, Sara Carucci, Lauren Z Atkinson, Luca Tessari, Tobias Banaschewski, David Coghill et al., "Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis," *The Lancet Psychiatry*, 2018, *5* (9), 727–738.

Cuddy, Emily and Janet Currie, "Rules vs. Discretion: Treatment of Mental Illness in U.S. Adolescents," Working Paper 27890, National Bureau of Economic Research October 2020.

Cullen, Julie Berry, "The impact of fiscal incentives on student disability rates," *Journal of Public Economics*, 2003, *87* (7-8), 1557–1589.

Currie, Janet and W Bentley Macleod, "Understanding Physician Decision Making: The Case of Depression," 2020.

_ , Mark Stabile, and Lauren Jones, "Do stimulant medications improve educational and behavioral outcomes for children with ADHD?," *Journal of health economics*, 2014, *37*, 58–69.

Dahl, Gordon B, Dan-Olof Rooth, and Anders Stenberg, "Family Spillovers in Field of Study," Working Paper 27618, National Bureau of Economic Research July 2020.

Dahl, Gordon B., Katrine V. LÃžken, and Magne Mogstad, "Peer Effects in Program Participation," *American Economic Review*, July 2014, *104* (7), 2049–74.

Dalsgaard, Søren, Maria Knoth Humlum, Helena Skyt Nielsen, and Marianne Simonsen, "Relative standards in ADHD diagnoses: the role of specialist behavior," *Economics Letters*, 2012, *117* (3), 663–665.

Danielson, Melissa L, Rebecca H Bitsko, Reem M Ghandour, Joseph R Holbrook, Michael D Kogan, and Stephen J Blumberg, "Prevalence of parent-reported ADHD diagnosis and associated treatment among US children and adolescents, 2016," *Journal of Clinical Child & Adolescent Psychology*, 2018, *47* (2), 199–212.

Depew, Briggs and Ozkan Eren, "Born on the wrong day? School entry age and juvenile crime," *Journal of Urban Economics*, 2016, *96*, 73–90.

Dhuey, Elizabeth, David Figlio, Krzysztof Karbownik, and Jeffrey Roth, "School starting age and cognitive development," *Journal of Policy Analysis and Management*, 2019, *38* (3), 538–578.

Dustan, Andrew, "Family networks and school choice," *Journal of Development Economics*, 2018, *134*, 372–391.

Einav, Liran, Amy Finkelstein, Tamar Oostrom, Abigail J Ostriker, and Heidi L Williams, "Screening and selection: The case of mammograms," Working Paper w26162, National Bureau of Economic Research 2019.

Elder, Todd E, "The importance of relative standards in ADHD diagnoses: evidence based on exact birth dates," *Journal of health economics*, 2010, *29* (5), 641–656.

_ **and Darren H Lubotsky**, "Kindergarten entrance age and childrenÂs achievement impacts of state policies, family background, and peers," *Journal of human Resources*, 2009, *44* (3), 641–683.

Evans, William N, Melinda S Morrill, and Stephen T Parente, "Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children," *Journal of health economics*, 2010, *29* (5), 657–673.

Faraone, Stephen V, Joseph Biederman, Wei J Chen, B Krifcher, K Keenan, C Moore, S Sprich, and MT Tsuang, "Segregation analysis of attention deficit hyperactivity disorder.," *Psychiatric Genetics*, 1992.

Fletcher, Jason, Nicole L Hair, and Barbara L Wolfe, "Am I my brother's keeper? Sibling spillover effects: The case of developmental disabilities and externalizing behavior," Working Paper w18279, National Bureau of Economic Research 2012.

Frank, Richard G and Thomas G McGuire, "Economics and mental health," in M. V. Pauly, T. G. McGuire, and P. P. Barros, eds., *Handbook of Health Economics*, Vol. 1, Elsevier, 2000, pp. 893–954.

Fredriksson, Peter and Björn Öckert, "Life-cycle effects of age at school start," *The Economic Journal*, 2014, *124* (579), 977–1004.

Furzer, Jill, "Diagnostic Errors in Child Mental Health: Assessing Treatment Selection and Its Long-Term Consequences," 2020. University of Toronto, unpublished manuscript.

_ **, Elizabeth Dhuey, and Audrey Laporte**, "ADHD Misidentification in School: Causes and Mitigators," Working Paper 200001, Canadian Centre for Health Economics 2020.

Goodman, Joshua, Michael Hurwitz, Christine Mulhern, and Jonathan Smith, "O brother, where start thou? Sibling spillovers in college enrollment," Working Paper w26502, National Bureau of Economic Research 2019.

Halldner, Linda, Annika Tillander, Cecilia Lundholm, Marcus Boman, Niklas Långström, Henrik Larsson, and Paul Lichtenstein, "Relative immaturity and ADHD: findings from nationwide registers, parent-and self-reports," *Journal of child psychology and psychiatry*, 2014, *55* (8), 897–904.

Hinshaw, Stephen P and Richard M Scheffler, *The ADHD explosion: Myths, medication, money, and today's push for performance*, Oxford University Press, 2014.

Hjörne, Eva, "Det har skett en stor förändring hemma," *Elevidentitet, föräldrars motstånd och ADHD i den svenska skolan. Utbildning & demokrati*, 2012, pp. 91–96.

Humphreys, Kathryn L, Timothy Eng, and Steve S Lee, "Stimulant medication and substance use outcomes: a meta-analysis," *JAMA psychiatry*, 2013, *70* (7), 740–749.

Hurwitz, Michael, Jonathan Smith, and Jessica S Howell, "Student age and the collegiate pathway," *Journal of Policy Analysis and Management*, 2015, *34* (1), 59–84.

Imbens, Guido W and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, 2008, *142* (2), 615–635.

Jensen, Peter S, "A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder," *Archives of general psychiatry*, 1999, *56* (12), 1073–1086.

Joensen, Juanna Schrøter and Helena Skyt Nielsen, "Spillovers in education choice," *Journal of Public Economics*, 2018, *157*, 158–183.

Karbownik, Krzysztof and Umut Özek, "Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design," Working Paper w26411, National Bureau of Economic Research 2019.

Kawaguchi, Daiji, "Actual age at school entry, educational outcomes, and earnings," *Journal of the Japanese and International Economies*, 2011, *25* (2), 64–80.

Krabbe, EE, ED Thoutenhoofd, M Conradi, SJ Pijl, and L Batstra, "Birth month as predictor of ADHD medication use in Dutch school classes," *European Journal of Special Needs Education*, 2014, *29* (4), 571–578.

Kubik, Jeffrey D, "Incentives for the identification and treatment of children with disabilities: the supplemental security income program," *Journal of Public Economics*, 1999, *73* (2), 187–215.

Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen, "School starting age and the crime-age profile," *The Economic Journal*, 2017, *127* (602), 1096–1118.

Landersø, Rasmus Kløve, Helena Skyt Nielsen, and Marianne Simonsen, "Effects of school starting age on the family," *Journal of Human Resources*, 2019, pp. 1117–9174R1.

Layton, Timothy J, Michael L Barnett, Tanner R Hicks, and Anupam B Jena, "Attention deficit–hyperactivity disorder and month of school enrollment," *New England Journal of Medicine*, 2018, *379* (22), 2122–2130.

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

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

Levy, F, DA Hay, M McStephen, C Wood, and I. Waldman, "Attention deficit hyperactivity disorder," *J Am Acad Child Adolesc Psychiatry*, 1997, *36* (6), 737–44.

Marquardt, Kelli, "Mis(sed) Diagnosis: Physician Decision Making and ADHD," 2020. University of Arizona, unpublished manuscript.

McEwan, Patrick J and Joseph S Shapiro, "The benefits of delayed primary school enrollment discontinuity estimates using exact birth dates," *Journal of human Resources*, 2008, *43* (1), 1–29.

Merten, Eva Charlotte, Jan Christopher Cwik, Jürgen Margraf, and Silvia Schneider, "Overdiagnosis of mental disorders in children and adolescents (in developed countries)," *Child and adolescent psychiatry and mental health*, 2017, *11* (1), 5.

Miller, Meghan, Erica D. Musser, Gregory S. Young, Brent Olson, Robert D. Steiner, and Joel T. Nigg, "Sibling Recurrence Risk and Cross-aggregation of Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder," *JAMA Pediatrics*, 02 2019, *173* (2), 147–152.

Molina, Brooke SG, Stephen P Hinshaw, L Eugene Arnold, James M Swanson, William E Pelham, Lily Hechtman, Betsy Hoza, Jeffery N Epstein, Timothy Wigal, Howard B Abikoff et al., "Adolescent substance use in the multimodal treatment study of attention-deficit/hyperactivity disorder (ADHD)(MTA) as a function of childhood ADHD, random assignment to childhood treatments, and subsequent medication," *Journal of the American Academy of Child & Adolescent Psychiatry*, 2013, *52* (3), 250–263.

Morrill, Melinda Sandler, "Special education financing and ADHD medications: A bitter pill to swallow," *Journal of Policy Analysis and Management*, 2018, *37* (2), 384–402.

Morrow, Richard L, E Jane Garland, James M Wright, Malcolm Maclure, Suzanne Taylor, and Colin R Dormuth, "Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children," *CMAJ*, 2012, *184* (7), 755–762.

Moses, Tally, "Being treated differently: Stigma experiences with family, peers, and school staff among adolescents with mental health disorders," *Social science & medicine*, 2010, *70* (7), 985–993.

Mullainathan, Sendhil and Ziad Obermeyer, "Does machine learning automate moral hazard and error?," *American Economic Review*, 2017, *107* (5), 476–80.

N., Melissa L. Danielson Rebecca H. Bitsko Joseph R. Holbrook Michael D. Kogan Reem M. Ghandour Ruth Perou Stephen J. Blumberg Visser Susanna, "Trends in the Parent-Report of Health Care Provider-Diagnosed and Medicated Attention-Deficit/Hyperactivity Disorder: United States, 2003-2011," *Journal of The American Academy of Child & Adolescent Psychiatry*, 2014.

Nicoletti, Cheti and Birgitta Rabe, "Sibling spillover effects in school achievement," *Journal of Applied Econometrics*, 2019, *34* (4), 482–501.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, 2019, *366* (6464), 447–453.

Ohan, Jeneva L, Troy AW Visser, Melanie C Strain, and Linda Allen, "Teachers' and education students' perceptions of and reactions to children with and without the diagnostic label "ADHD"," *Journal of School Psychology*, 2011, *49* (1), 81–105.

Ong, Mei-Sing and Kenneth D Mandl, "National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at $4 billion a year," *Health Affairs*, 2015, *34* (4), 576–583.

Parman, John, "Childhood health and sibling outcomes: Nurture Reinforcing nature during the 1918 influenza pandemic," *Explorations in Economic History*, 2015, *58* (C), 22–43.

Pottegård, Anton, Jesper Hallas, and Helga Zoëga, "Children's relative age in class and use of medication for ADHD: a Danish Nationwide Study," *Journal of Child Psychology and Psychiatry*, 2014, *55* (11), 1244–1250.

Quadagno, Jill, "Incentives to disability in federal disability insurance and supplemental security income," *Clinical Orthopaedics and Related Research®*, 1997, *336*, 11–17.

Qureshi, Javaeria A., "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," *The Economic Journal*, 2018, *128* (616), 3285–3319.

Qureshi, Javaeria A, "Siblings, teachers, and spillovers on academic achievement," *Journal of Human Resources*, 2018, *53* (1), 272–297.

Root, Adrian, Jeremy P Brown, Harriet J Forbes, Krishnan Bhaskaran, Joseph Hayes, Liam Smeeth, and Ian J Douglas, "Association of relative age in the school year with diagnosis of intellectual disability, attention-deficit/hyperactivity disorder, and depression," *JAMA pediatrics*, 2019, *173* (11), 1068–1075.

Schwandt, Hannes and Amelie Wuppermann, "The youngest get the pill: ADHD misdiagnosis in Germany, its regional correlates and international comparison," *Labour Economics*, 2016, *43*, 72–86.

Schwarz, Alan, *ADHD nation: Children, doctors, big pharma, and the making of an American epidemic*, Simon and Schuster, 2017.

Socialstyrelsen, "Socialstyrelsens riktlinjer för skolhälsovården," *Socialstyrelsens Rapporter*, 2002.

_ , "Användning av centralstimulantia vid adhd," *Socialstyrelsens Rapporter*, 2012.

_ , "Utredning och diagnostik av adhd hos barn och ungdomar," *Socialstyrelsen*, 2014.

_ , "Förskrivning av centralstimulerande läkemedel vid adhd," *Socialstyrelsens Rapporter*, 2015.

Tarver, Joanne, David Daley, and Kapil Sayal, "Attention-deficit hyperactivity disorder (ADHD): an updated review of the essential facts," *Child: care, health and development*, 2014, *40* (6), 762–774.

Thapar, Anita and Miriam Cooper, "Attention deficit hyperactivity disorder," *The Lancet*, 2016, *387* (10024), 1240–50.

Thomas, Rae, Sharon Sanders, Jenny Doust, Elaine Beller, and Paul Glasziou, "Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis," *Pediatrics*, 2015, *135* (4), e994–e1001.

Whitely, Martin, Melissa Raven, Sami Timimi, Jon Jureidini, John Phillimore, Jonathan Leo, Joanna Moncrieff, and Patrick Landman, "Attention deficit hyperactivity disorder late birthdate effect common in both high and low prescribing international jurisdictions: systematic review.," *Journal of Child Psychology and Psychiatry*, 2018.

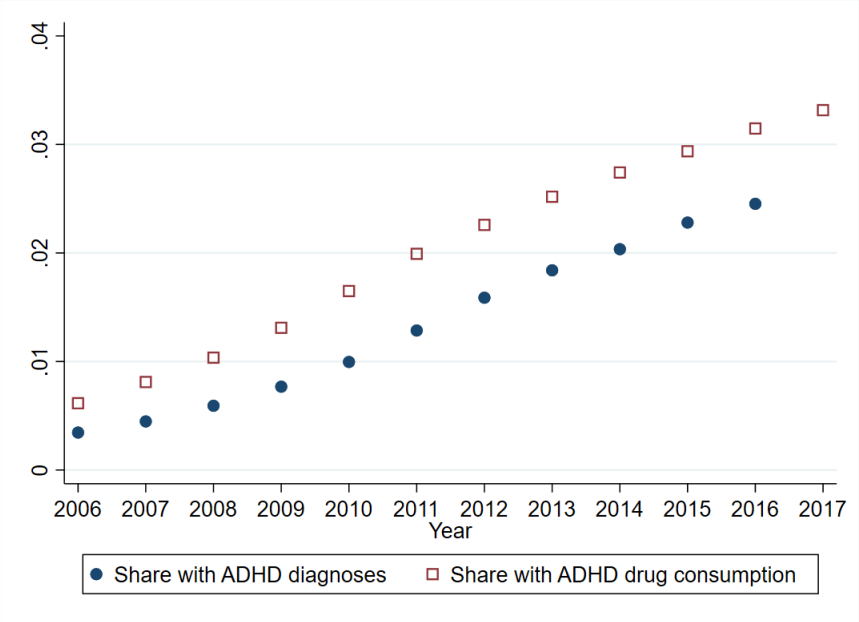
Wilens, Timothy E, Stephen V Faraone, Joseph Biederman, and Samantha Gunawardene, "Does stimulant therapy of attention-deficit/hyperactivity disorder beget later substance abuse? A meta-analytic review of the literature," *Pediatrics*, 2003, *111* (1), 179–185.

Yi, Junjian, James J Heckman, Junsen Zhang, and Gabriella Conti, "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes," *The Economic Journal*, 2015, *125* (588), F347–F371.

Zoëga, Helga, Unnur A Valdimarsdóttir, and Sonia Hernández-Díaz, "Age, academic performance, and stimulant prescribing for ADHD: a nationwide cohort study," *Pediatrics*, 2012, *130* (6), 1012–1018.

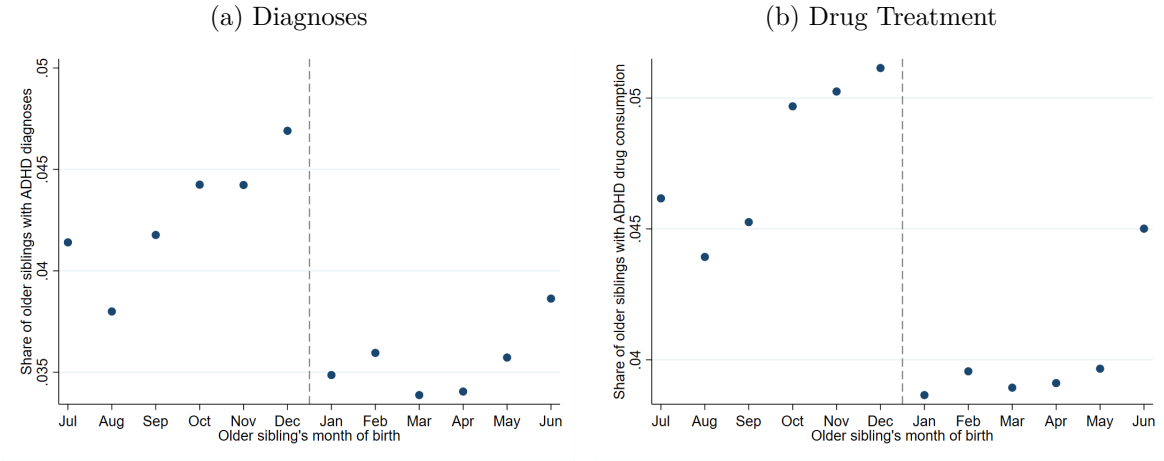
7 Figures and Tables

Figure 1: Trends in the Share of Children Ages 6–19 with ADHD



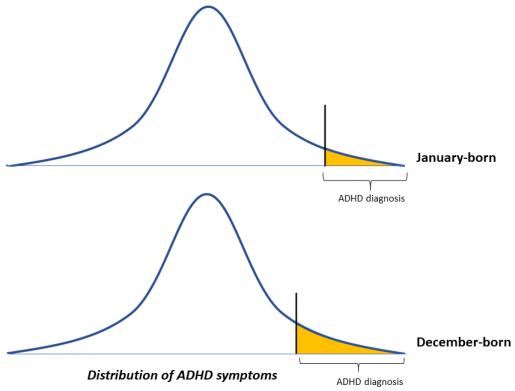
Note: The sample includes children between the ages of 6 and 19 who are born in Sweden, Norway, or the EU 28. For every year, the figure plots the share of these children with at least one ADHD diagnosis in the outpatient data (in blue-filled dots) and at least one ADHD drug claim (in red-outlined squares), respectively.

Figure 2: ADHD Diagnoses and Drug Treatment by Own Month of Birth, Older Siblings Only



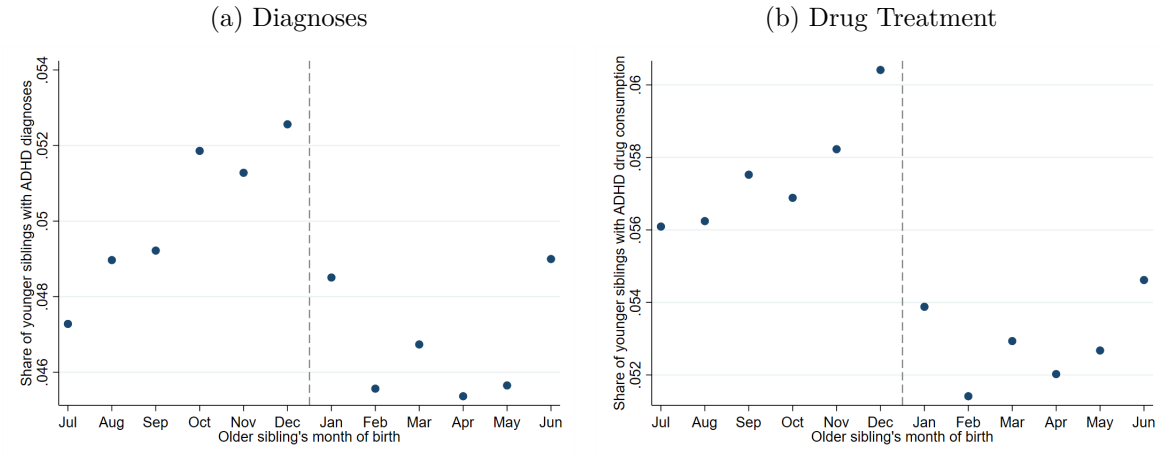
Note: The sample of analysis is the universe of singleton sibling pairs born in Sweden, Norway, or the EU 28, where the older sibling is born between January 1990 and December 1996. These graphs plot ADHD-related outcomes for older siblings by their own birth month. Sub-figure (a) plots the share of older siblings with an ADHD diagnosis in the outpatient data, while sub-figure (b) plots the share of older siblings with at least one ADHD drug claim in the prescription drug data.

Figure 3: Stylized Framework for Interpreting ADHD Gap



Note: This figure depicts a stylized visual framework for interpreting the ADHD gap at the school-entry cutoff. The bell curves represent the distributions of underlying ADHD symptoms in the populations of children born in December and January, respectively. The yellow areas under each of the curves signify the children who receive a positive ADHD diagnosis.

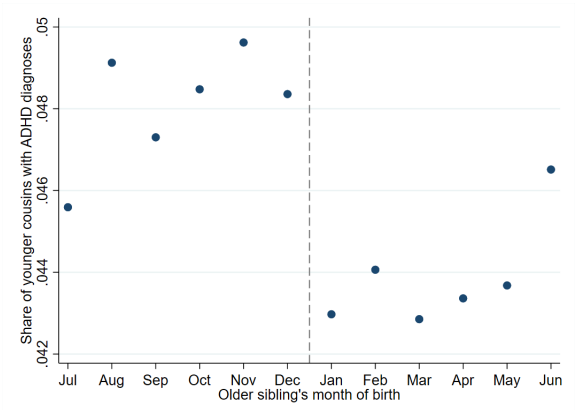
Figure 4: Younger Siblings' ADHD Diagnoses and Drug Treatment by Older Sibling's Month of Birth



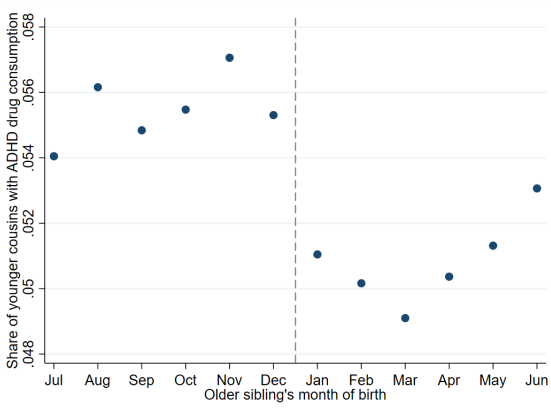
Note: The sample of analysis is the universe of singleton sibling pairs born in Sweden, Norway, or the EU 28, where the older sibling is born between January 1990 and December 1996. These graphs plot ADHD-related outcomes for younger siblings (on the y-axes) by the birth month of the older sibling (on the x-axes). Sub-figure (a) plots the share of younger siblings with an ADHD diagnosis in the outpatient data, while sub-figure (b) plots the share of younger siblings with at least one ADHD drug claim in the prescription drug data.

Figure 5: Younger Cousins' ADHD Diagnoses and Drug Treatment by Older Cousin's Month of Birth

(a) Diagnoses



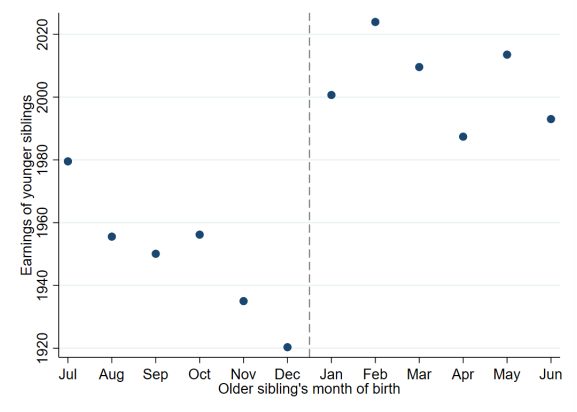
(b) Drug Treatment



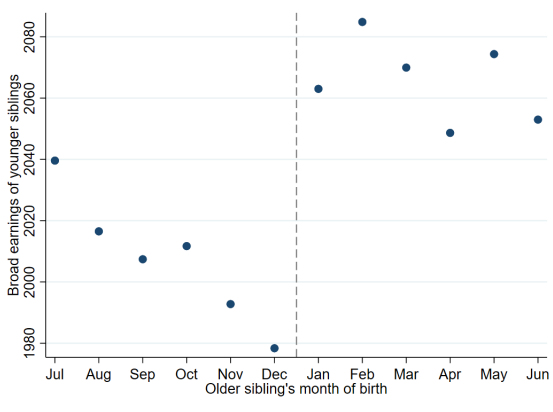
Note: To construct the analysis sample for these figures, we identify all cousins of the older siblings in our main sample (see notes under Figure 2), and then keep for each older sibling, his/her next younger cousin. We only keep cousins who are born in Sweden, Norway, or the EU 28. These figures plot ADHD-related outcomes of the younger cousins (on the y-axes) by the birth month of the older cousin (on the x-axes). Sub-figure (a) plots the share of younger cousins with an ADHD diagnosis in the outpatient data, while sub-figure (b) plots the share of younger cousins with at least one ADHD drug claim in the prescription drug data.

Figure 6: Younger Siblings' Educational and Labor Market Outcomes by Older Sibling's Month of Birth

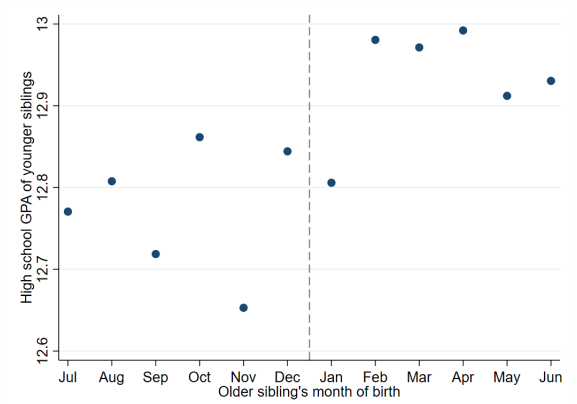
(a) Earnings, 2018



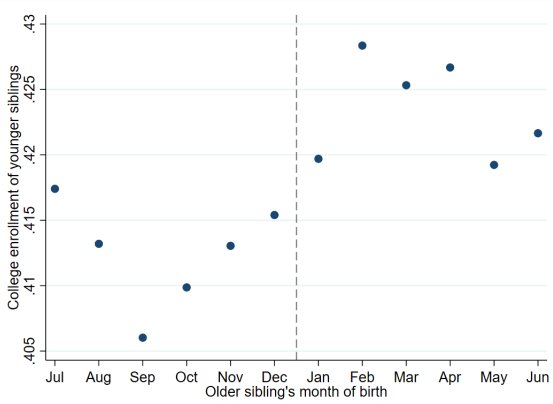
(b) Broad Earnings, 2018



(c) HS GPA, 2016



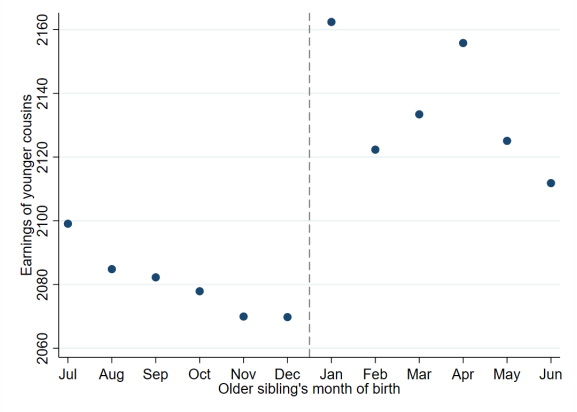
(d) College Enrollment, by 2018



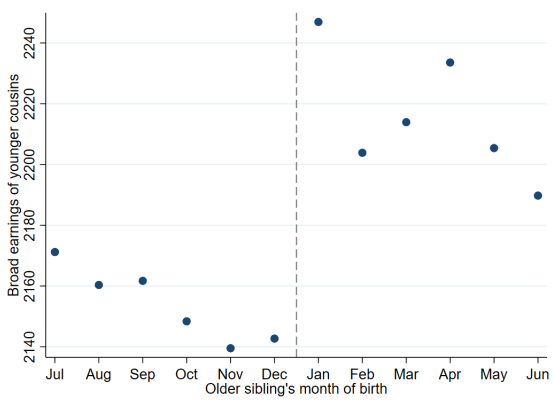
Note: These figures plot average educational and labor market outcomes of younger siblings (on the y-axes), by the birth month of the older sibling (on the x-axes). The sample is limited to sibling pairs with younger siblings born in 1990-1995 only (see notes under Figure 2 for further description of the main siblings sample).

Figure 7: Younger Cousins' Educational and Labor Market Outcomes by Older Cousin's Month of Birth

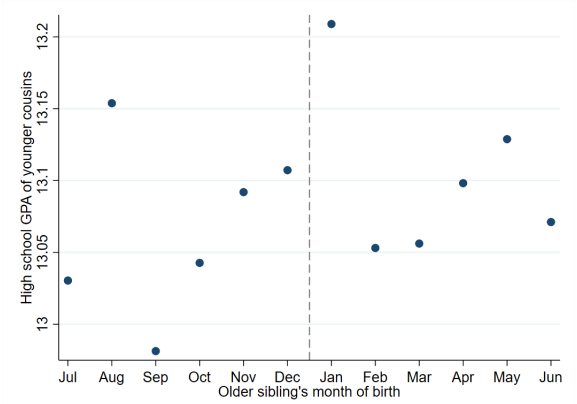
(a) Earnings, 2018



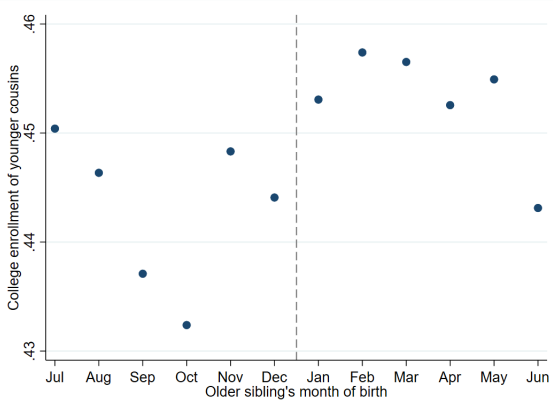
(b) Broad Earnings, 2018



(c) HS GPA, 2016



(d) College Enrollment, by 2018



Note: These figures plot average educational and labor market outcomes of younger cousins (on the y-axes), by the birth month of the older cousin (on the x-axes). The sample is limited to cousin pairs with younger cousins born in 1990-1995 only (see notes under Figure 5 for further description of the main cousins sample).

Table 1: Effect of Older Sibling Being Born After Cutoff on Own ADHD Diagnosis and Drug Treatment

	ADHD Diagnoses		ADHD Drugs		
	(1) Main Sample	(2) Age Diff ≤ 3	(3) Main Sample	(4) Age Diff ≤ 3	(5) Balanced Timing
Born after cutoff	-0.0117***	-0.0122***	-0.0125***	-0.0127***	-0.0112***
[Own Relative Age Effect]	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Mean(Y)	0.039	0.040	0.044	0.045	0.025
N	271,388	175,445	271,388	175,445	271,388

Notes: Each column reports results from a separate regression estimating model (1). The sample of analysis is the universe of older siblings born between January 1990 and December 1996, among singleton sibling pairs born in Sweden, Norway, or the EU 28. The dependent variable in columns (1) and (2) is an indicator equal to one if the older sibling ever has an outpatient claim with an ADHD diagnosis. The dependent variable in columns (3) and (4) is an indicator equal to one if the older sibling has at least one ADHD drug claim in the prescription drug data. In columns (2) and (4), the main sample is further limited to sibling pairs with an age difference of no more than 3 years (36 months). The dependent variable in column (5) is an indicator equal to one if the older sibling has at least one ADHD drug claim that is specifically recommended for children between ages 5 and 18. In addition, the observation period for drug claims in column (5) is restricted such that children born in the same fiscal year (July-June) are observed for the same number of months (see Appendix Table B1). All regressions control for a linear spline function for the older sibling's month of birth centered around January (i.e., the running variable in the RD specification), and also include controls for an indicator for whether the older sibling is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth (high school only, some college, college degree or more), the log household income averaged over the year of the older child's birth and the following two years, and fixed effects for the fiscal years of birth of the older siblings. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 2: Effect of Older Sibling Being Born After Cutoff on Younger Sibling's ADHD Diagnosis and Drug Treatment

	ADHD Diagnoses		ADHD Drugs		
	(1) Main Sample	(2) Age Diff ≤ 3	(3) Main Sample	(4) Age Diff ≤ 3	(5) Balanced Timing
Older sibling born after cutoff	-0.0059***	-0.0072***	-0.0055***	-0.0068***	-0.0046***
[Spillover Effect]	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Younger sibling born after cutoff	-0.0155***	-0.0148***	-0.0164***	-0.0155***	-0.0151***
[Own Relative Age Effect]	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Mean(Y)	0.049	0.048	0.056	0.054	0.043
N	271,388	175,445	271,388	175,445	271,388

Notes: Each column reports results from a separate regression estimating model (2). The sample of analysis is the universe of singleton sibling pairs born in Sweden, Norway, or the EU 28, where the older sibling is born between January 1990 and December 1996. The dependent variable in columns (1) and (2) is an indicator equal to one if the younger sibling ever has an outpatient claim with an ADHD diagnosis. The dependent variable in columns (3) and (4) is an indicator equal to one if the younger sibling has at least one ADHD drug claim in the prescription drug data. In columns (2) and (4), the main sample is further limited to sibling pairs with an age difference of no more than 3 years (36 months). The dependent variable in column (5) is an indicator equal to one if the younger sibling has at least one ADHD drug claim that is specifically recommended for children between ages 5 and 18. In addition, the observation period for drug claims in column (5) is restricted such that children born in the same fiscal year (July-June) are observed for the same number of months (see Appendix Table B1). All regressions control for linear spline functions of the older and younger child's month of birth centered around January (i.e., the running variables in the RD specifications), and also include controls for birth spacing (in months), indicators for whether the older and younger sibling is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth (high school only, some college, college degree or more), the log household income averaged over the year of the older child's birth and the following two years, and fixed effects for the fiscal years of birth of the older and younger siblings. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 3: Effect of Older Cousin Being Born After Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment

	ADHD Diagnoses		ADHD Drugs		
	(1) Main Sample	(2) Age Diff ≤ 3	(3) Main Sample	(4) Age Diff ≤ 3	(5) Balanced Timing
Older cousin born after cutoff	-0.0060***	-0.0097***	-0.0048***	-0.0089***	-0.0050***
[Spillover Effect]	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Younger cousin born after cutoff	-0.0187***	-0.0158***	-0.0198***	-0.0148***	-0.0195***
[Own Relative Age Effect]	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mean(Y)	0.046	0.044	0.053	0.049	0.040
N	208,437	126,380	208,437	126,380	208,437

Notes: To construct the analysis sample for these regressions, we identify all cousins of the older siblings in our main sample, and then keep for each older sibling, his/her next younger cousin. We only keep cousins who are born in Sweden, Norway, or the EU 28. See notes under Table 2 for more details about the sample, specifications, control variables, and outcomes. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 4: Effect of Older Sibling/Cousin Being Born After Cutoff on Younger Siblings' and Cousins' Educational and Labor Market Outcomes

	(1) Earnings	(2) Broad Earnings	(3) HS GPA	(4) Coll Enroll
Panel A: Younger Siblings				
Older sibling born after cutoff	32.8370***	29.5572***	0.1336	0.0086***
[Spillover Effect]	(7.846)	(7.100)	(0.089)	(0.003)
Younger sibling born after cutoff	-75.1359***	-80.9148***	0.4663***	-0.0144***
[Own Relative Age Effect]	(22.487)	(22.250)	(0.054)	(0.005)
Mean(Y)	2001.360	2061.876	12.948	0.420
N	142,034	142,034	125,824	144,294
Panel B: Younger Cousins				
Older cousin born after cutoff	10.8254	9.3314	0.1309**	0.0120*
[Spillover Effect]	(13.527)	(14.385)	(0.049)	(0.006)
Younger cousin born after cutoff	-65.0976***	-75.8078***	0.6099***	0.0058
[Own Relative Age Effect]	(14.320)	(14.052)	(0.081)	(0.005)
Mean(Y)	2116.893	2194.843	13.105	0.450
N	108,588	108,588	96,880	110,561

Notes: Each column in each panel reports results from a separate regression, estimating equation (2). We restrict our main sample to sibling and cousin pairs where the younger sibling/cousin is born in 1990-1995 only. The earnings measures are measured in 2018. High school GPA is measured in 2016. College enrollment is an indicator set to 1 if an individual is ever enrolled in college by 2018. "Broad earnings" refers to income from employment and employment-based transfers such as parental leave benefits. See notes under Table 2 for more details about the sample, specifications, and control variables. Robust standard errors are clustered on the older child's month of birth.

Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

ONLINE APPENDIX

A Symptoms and Diagnosis of ADHD

Health care providers use the guidelines in the American Psychiatric Association's Diagnostic and Statistical Manual, Fifth edition (DSM-5) to diagnose ADHD.[40] Individuals with ADHD show a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development. The following are listed as symptoms of ADHD:

Inattention Symptoms:

1. Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or with other activities.
2. Often has trouble holding attention on tasks or play activities.
3. Often does not seem to listen when spoken to directly.
4. Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., loses focus, side-tracked).
5. Often has trouble organizing tasks and activities.
6. Often avoids, dislikes, or is reluctant to do tasks that require mental effort over a long period of time (such as schoolwork or homework).
7. Often loses things necessary for tasks and activities (e.g. school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones).
8. Is often easily distracted.
9. Is often forgetful in daily activities.

Hyperactivity and Impulsivity Symptoms:

1. Often fidgets with or taps hands or feet, or squirms in seat.
2. Often leaves seat in situations when remaining seated is expected.

[40] `https://www.cdc.gov/ncbddd/adhd/diagnosis.html`

3. Often runs about or climbs in situations where it is not appropriate (adolescents or adults may be limited to feeling restless).
4. Often unable to play or take part in leisure activities quietly.
5. Is often "on the go" acting as if "driven by a motor".
6. Often talks excessively.
7. Often blurts out an answer before a question has been completed.
8. Often has trouble waiting their turn.
9. Often interrupts or intrudes on others (e.g., butts into conversations or games).

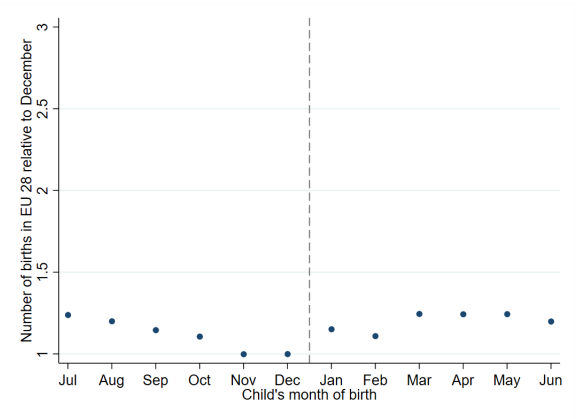
An ADHD diagnosis is indicated when the following conditions must be met:

- Six or more symptoms of inattention for children up to age 16 years, or five or more symptoms for individuals age 17 years and older.
- Symptoms have been present for at least 6 months to an extent that is disruptive or inappropriate for the person's developmental level.
- Several inattentive or hyperactive-impulsive symptoms were present before age 12 years.
- Several symptoms are present in two or more settings (such as at home, school or work; with friends or relatives; in other activities).
- There is clear evidence that the symptoms interfere with, or reduce the quality of, social, school, or work functioning.
- The symptoms are not better explained by another mental disorder (such as a mood disorder, anxiety disorder, dissociative disorder, or a personality disorder). The symptoms do not happen only during the course of schizophrenia or another psychotic disorder.

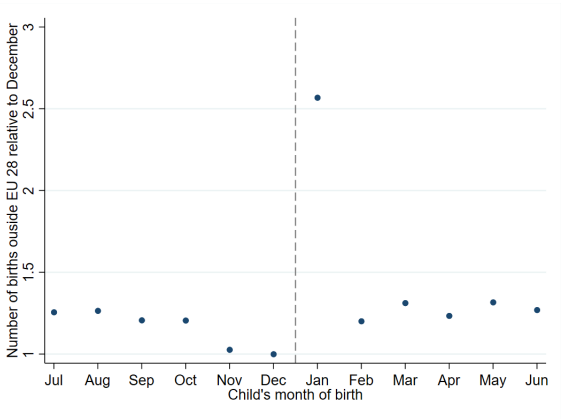
B Additional Results

Figure B1: Number of Births by Month of Birth, January 1990 to December 2016 Births

(a) Number of Births in Sweden / EU 28

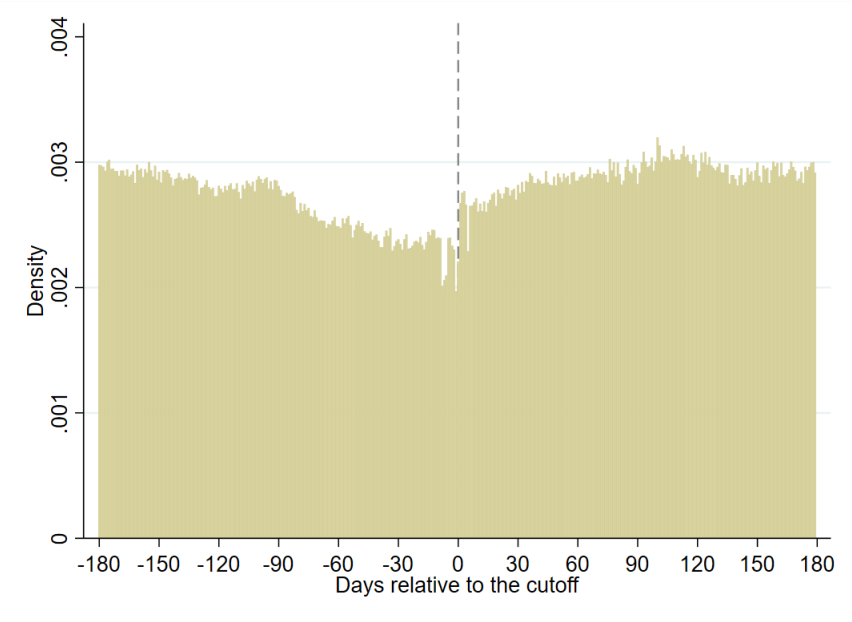


(b) Number of Births Outside EU 28



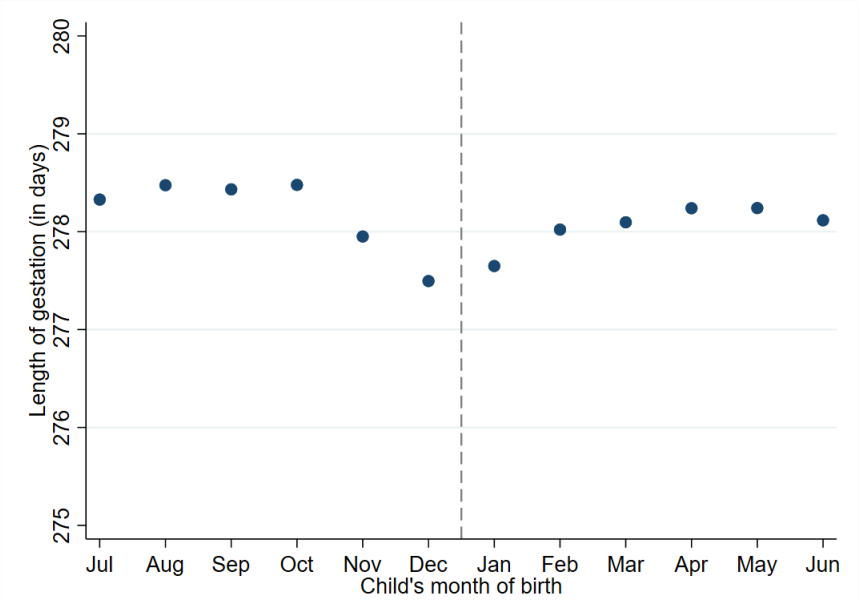
Note: The figures plot the number of births as a ratio relative to the number of births in December, by birth month for Swedish residents born in Sweden, Norway, and EU 28 (sub-figure a) and Swedish residents born in the rest of the world (sub-figure b), using all births between January 1990 and December 2016.

Figure B2: Distribution of Births at the Daily Level, Children Born Between January 1995 and December 2014



Note: The sample includes all children born in Sweden between January 1995 and December 2014 with information on exact date of birth. The figure shows a histogram of the distribution of births at the daily level, with a bandwidth of 180 days around the cutoff (January 1).

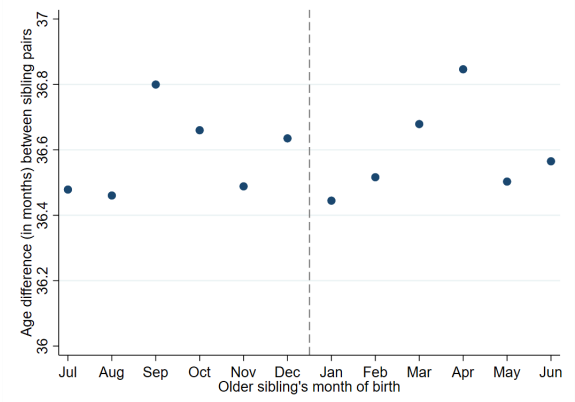
Figure B3: Average Gestation Length by Month of Birth, Children Born Between January 1995 and December 2014



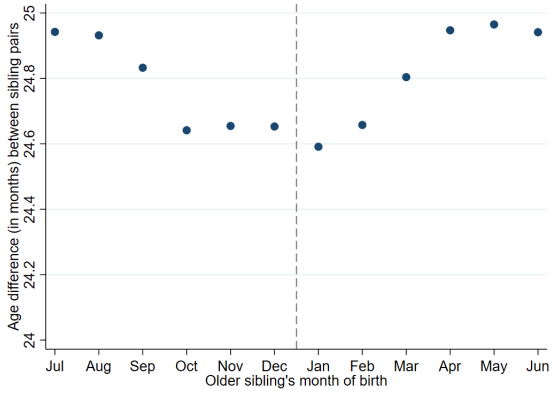
Note: The sample includes all children born in Sweden between January 1995 and December 2014 with information on gestation length. The figure plots the average length of gestation in days by the child's birth month.

Figure B4: Age Difference Between Siblings by Month of Birth of Older Sibling

(a) All Sibling Pairs



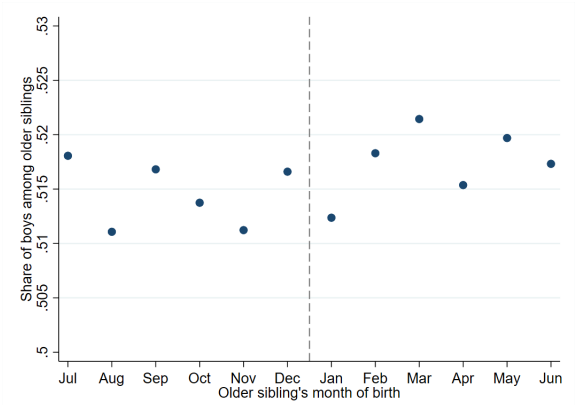
(b) Age Difference $\leq$ 3 Years



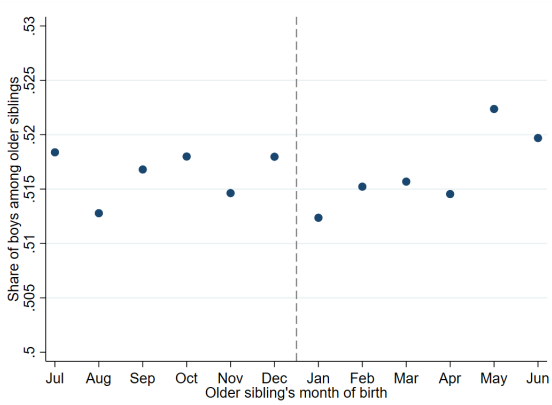
Note: See notes under Figure 2 for more information about the sample. These figures plot the average age difference between siblings (in months) by the birth month of the older sibling. Sub-figure (a) uses our main sample of sibling pairs, while sub-figure (b) limits our main sample to the pairs with an age difference of no more than 3 years (36 months).

Figure B5: Sibling Gender Composition by Month of Birth of Older Sibling

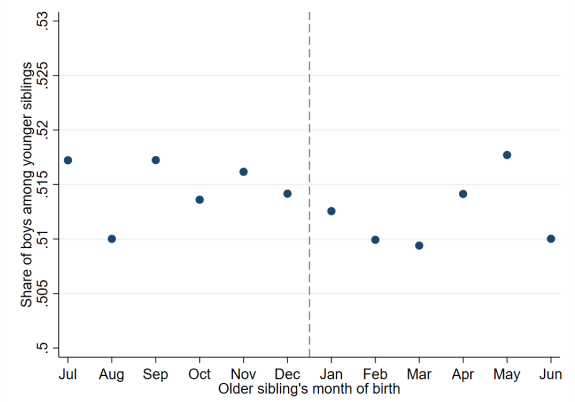
(a) Share Boys, Older Siblings



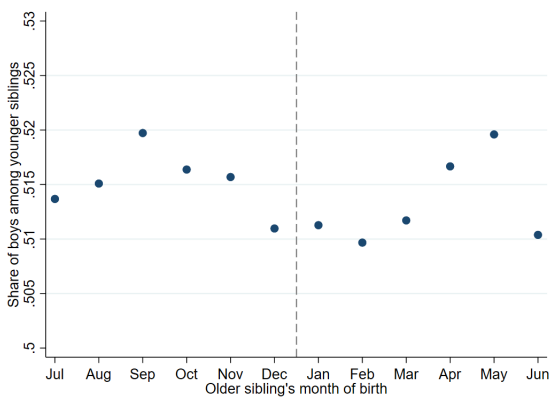
(b) Share Boys, OS, Age Diff ≤ 3



(c) Share Boys, Younger Siblings



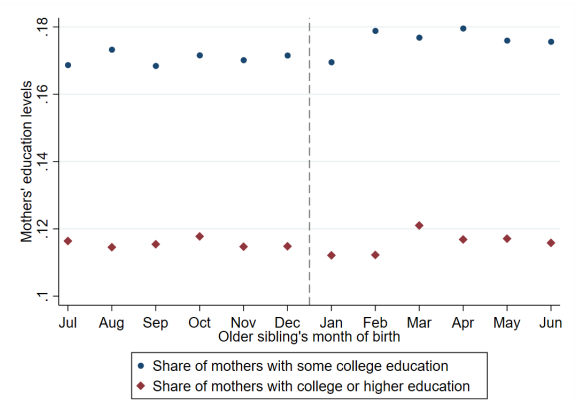
(d) Share Boys, YS, Age Diff ≤ 3



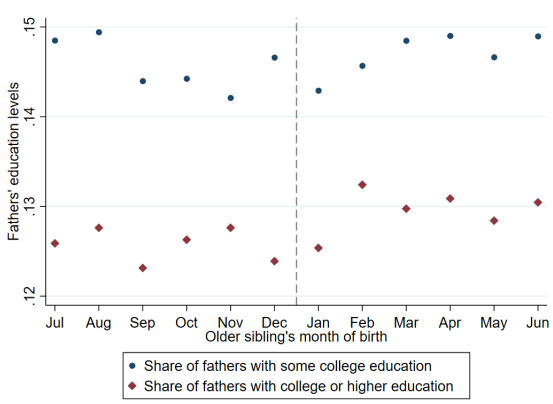
Note: See notes under Figure 2 for more information about the sample. These figures plot the share of boys among older and younger siblings by the birth month of the older sibling. Sub-figure (a) plots the share of boys among older siblings in our main sample of sibling pairs, and sub-figure (b) limits our main sample to pairs with an age difference of no more than 3 years (36 months). Sub-figure (c) plots the share of boys among younger siblings in our main sample of sibling pairs, and sub-figure (d) limits our main sample to pairs with an age difference of no more than 3 years (36 months).

Figure B6: Parental Education Level by Month of Birth of Older Child

(a) Mothers



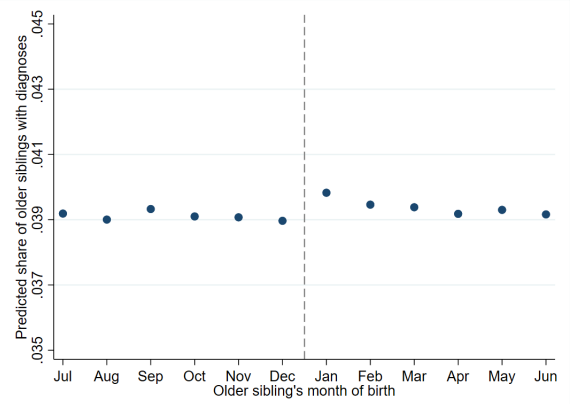
(b) Fathers



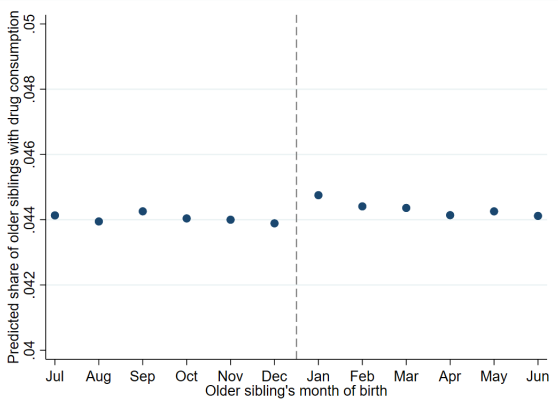
Note: See notes under Figure 2 for more information about the sample. Sub-figure (a) plots the share of mothers with some college education (in blue) and the share of mothers with college or higher level of education (in red), by their older child's month of birth. Sub-figure (b) plots the share of fathers with some college education (in blue) and the share of fathers with college or higher level of education (in red), by their older child's month of birth.

Figure B7: Predicted ADHD Outcomes by Month of Birth of Older Sibling

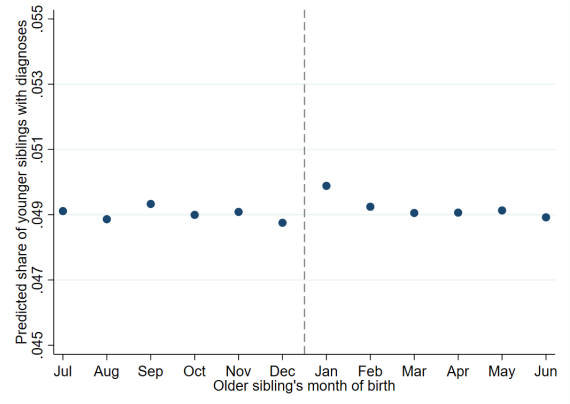
(a) Pred. Diagnoses, Older Sibs



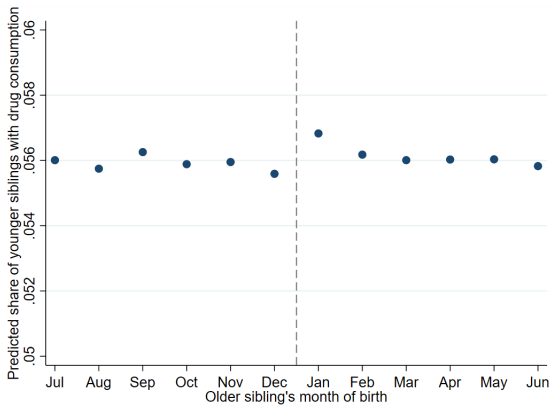
(b) Pred. Drug Treatment, Older Sibs



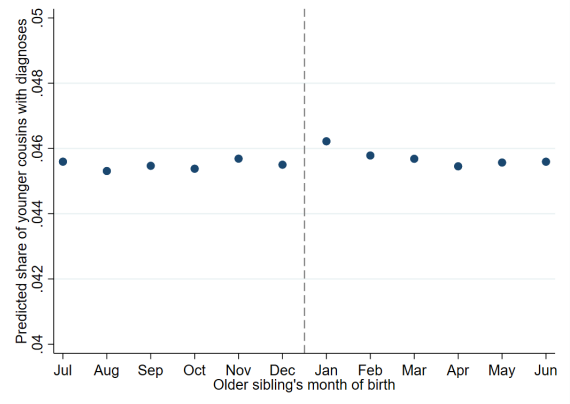
(c) Pred. Diagnoses, Younger Sibs



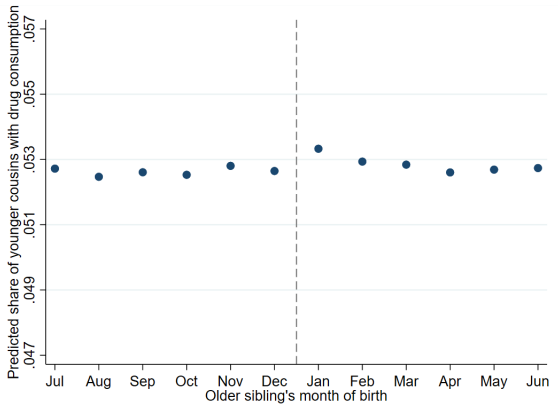
(d) Pred. Drug Treatment, Younger Sibs



(e) Pred. Diagnoses, Younger Cousins

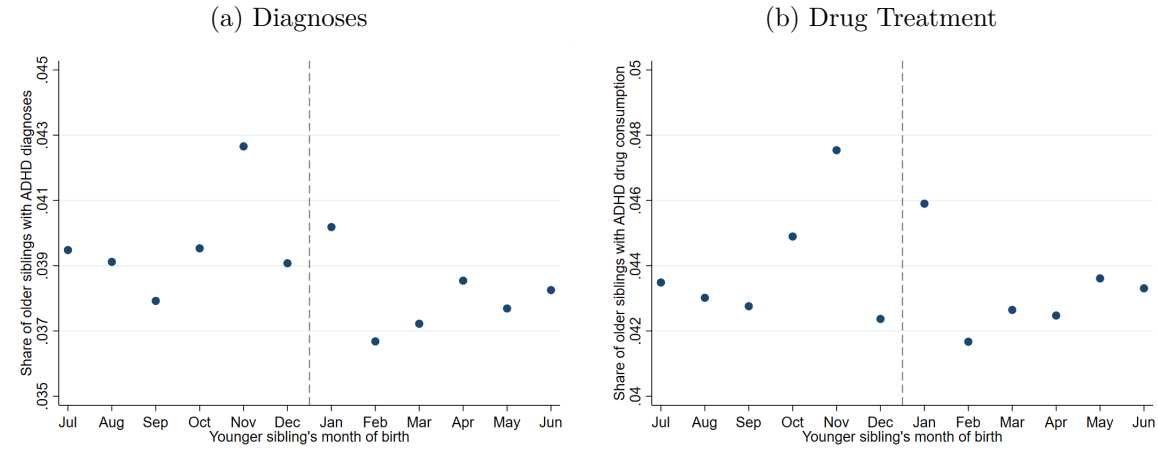


(f) Pred. Drug Treatment, Younger Cousins



Note: See notes under Figure 2 for more information about the sample. These graphs plot predicted ADHD-related outcomes for the older siblings, younger siblings, and younger cousins, respectively, by the birth month of the older sibling. The predicted outcomes are constructed by regressing each ADHD outcome on birth spacing (in months), indicators for whether the older and younger sibling is male, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older child's birth (high school only, some college, college degree or more), and the log household income averaged over the year of the older child's birth and the following two years.

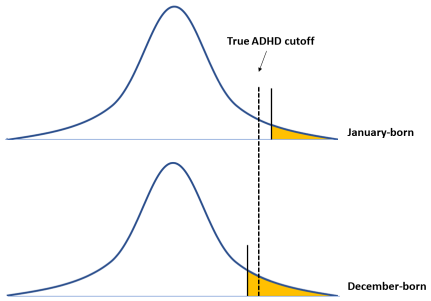
Figure B8: Older Sibling's ADHD Diagnoses and Drug Treatment by Younger Sibling's Month of Birth



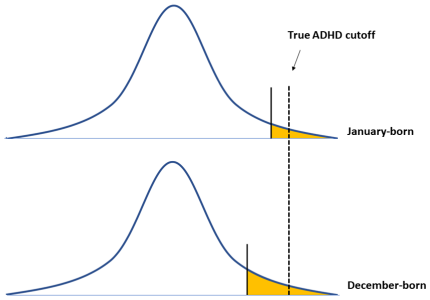
Note: See notes under Figure 2 for more information about the sample. These figures plot the share of older siblings who ever have an ADHD diagnosis in the outpatient data (subfigure a) and who have at least one ADHD drug claim in the prescription drug data (subfigure b) by the birth month of the younger sibling.

Figure B9: Stylized Framework for Interpreting ADHD Gap; Three Possible Interpretations

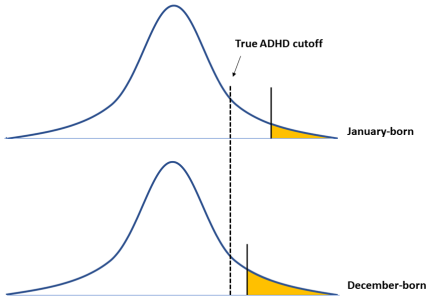
(a) Over-Diagnosis in December, Under-Diagnosis in January



(b) More Over-Diagnosis in December than January



(c) Less Under-Diagnosis in December than January



Note: These figures depict a stylized visual framework for interpreting the ADHD gap at the school-entry cutoff. The bell curves represent the distributions of underlying ADHD risk in the populations of children born in December and January, respectively. The yellow areas under each of the curves signify the children who receive a positive ADHD diagnosis. The vertical dashed line in each sub-figure represents different assumptions about the underlying "natural rate" of ADHD in the population, which is assumed to be independent of the child's birth month.

Table B1: Balanced Timing Drug Claim Observation Period

Month/Year of Birth	Drug Claim Observation Period	Ages Observed
01/1990 - 06/1990	07/2005 - 06/2008	16y0m - 18y0m
07/1990 - 06/1991	07/2005 - 06/2009	15y0m - 18y0m
07/1991 - 06/1992	07/2005 - 06/2010	14y0m - 18y0m
07/1992 - 06/1993	07/2005 - 06/2011	13y0m - 18y0m
07/1993 - 06/1994	07/2005 - 06/2012	12y0m - 18y0m
07/1994 - 06/1995	07/2005 - 06/2013	11y0m - 18y0m
07/1995 - 06/1996	07/2005 - 06/2014	10y0m - 18y0m
07/1996 - 06/1997	07/2005 - 06/2015	9y0m - 18y0m
07/1997 - 06/1998	07/2005 - 06/2016	8y0m - 18y0m
07/1998 - 06/1999	07/2005 - 06/2017	7y0m - 18y0m
07/1999 - 06/2000	07/2005 - 06/2017	6y0m - 17y0m
07/2000 - 06/2001	07/2005 - 06/2017	5y0m - 16y0m
07/2001 - 06/2002	07/2006 - 06/2017	5y0m - 15y0m
07/2002 - 06/2003	07/2007 - 06/2017	5y0m - 14y0m
07/2003 - 06/2004	07/2008 - 06/2017	5y0m - 13y0m
07/2004 - 06/2005	07/2009 - 06/2017	5y0m - 12y0m
07/2005 - 06/2006	07/2010 - 06/2017	5y0m - 11y0m
07/2006 - 06/2007	07/2011 - 06/2017	5y0m - 10y0m
07/2007 - 06/2008	07/2012 - 06/2017	5y0m - 9y0m
07/2008 - 06/2009	07/2013 - 06/2017	5y0m - 8y0m
07/2009 - 06/2010	07/2014 - 06/2017	5y0m - 7y0m

Notes: This table reports the drug claim observation period for all individuals in our data when we use the "balanced timing" definition of ADHD drug claims. Specifically, our aim is to measure ADHD drug claims over the same number of months for all children within each fiscal year (July-June) of birth.

Table B2: Sample Means of Key Variables

	Full Sample	Jul-Dec Births	Jan-Jun Births
Older siblings, share w/ ADHD diagnosis	0.039	0.043	0.035
Older siblings, share w/ ADHD drug use	0.044	0.048	0.040
Younger siblings, share w/ ADHD diagnosis	0.048	0.050	0.047
Younger siblings, share w/ ADHD drug use	0.055	0.057	0.053
Younger cousins, share w/ ADHD diagnosis	0.046	0.048	0.044
Younger cousins, share w/ ADHD drug use	0.053	0.055	0.051
Birth spacing (in months)	36.591	36.585	36.597
Mother is foreign-born	0.153	0.162	0.146
Father is foreign-born	0.165	0.173	0.157
Log household income	7.493	7.487	7.498
Mother has college degree+	0.116	0.116	0.116
Father has college degree+	0.128	0.126	0.130
Observations	293,994	136,308	157,686

Notes: This table reports sample means of some of the variables in our analysis. The first column uses our full analysis sample of singleton sibling pairs born in Sweden, Norway, or the EU 28, where the older sibling is born between January 1990 and December 1996. The second and third columns split the sample into families with older siblings born in July-December and January-June, respectively.

Table B3: Results for Placebo Outcomes

	(1) Gestation length	(2) Birth spacing	(3) Share boys (OS)	(4) Share boys (YS)	(5) Fathers' educ	(6) Mothers' educ
Born after cutoff	0.1645					
	(0.291)					
Older sibling born after cutoff		-0.0241	0.0015	-0.0008	0.0056*	0.0040*
		(0.017)	(0.004)	(0.002)	(0.003)	(0.002)
Younger sibling born after cutoff		6.1786***	-0.0062	-0.0005	-0.0068***	-0.0042*
		(0.011)	(0.004)	(0.003)	(0.002)	(0.002)
			(0.001)	(0.001)	(0.001)	(0.000)
Mean(Y)	278.155	36.289	0.516	0.513	0.129	0.117
N	1,891,751	271,388	271,388	271,388	271,388	271,388

Notes: Each column reports results from a separate regression. In column (1), the sample includes children born between January 1995 and December 2014 with information on gestation length, while we use our main sample of sibling pairs in columns (2)-(6) (see notes under Table 2). The dependent variables in columns (5) and (6) are indicators for whether each parent has a college degree or higher. In column (1), we control for a linear spline in the running variable (month of birth centered around January), as well as the following control variables: indicator for the child being male, indicators for each parent being foreign-born, indicators for each parent's education category (high school only, some college, college degree or more), the log household income averaged over the year of birth and the following two years, and fiscal year of birth fixed effects. In columns (2)-(6), we include the same control variables as in our main specifications (see notes under Table 2), except we omit birth spacing in column (2) and indicators for the older and younger sibling's child gender in columns (3) and (4), respectively. We also omit the father's education categories in column (5) and the mother's education categories in column (6). In column (2), we also drop the younger sibling's birth month and the interaction term with being born after the cutoff due to collinearity. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B4: Effect of Older Sibling Being Born After Cutoff on Own ADHD Diagnoses and Drug Treatment, Controlling for Younger Sibling's Relative Age

	Diagnoses		Drugs		
	(1) Main Sample	(2) Age Diff ≤ 3	(3) Main Sample	(4) Age Diff ≤ 3	(5) Balanced Timing
Older sibling born after cutoff	-0.0116***	-0.0118***	-0.0124***	-0.0125***	-0.0113***
[Own Relative Age Effect]	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Younger sibling born after cutoff	-0.0022	-0.0033**	-0.0009	-0.0002	-0.0016
[Younger Sib. Relative Age Effect]	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Mean(Y)	0.039	0.040	0.044	0.045	0.025
N	271,388	175,445	271,388	175,445	271,388

Notes: Each column reports results from a separate regression. See notes under Table 2 for more details about the sample, specifications, control variables, and outcomes. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * p < 0.1 ** p < 0.05 *** p < 0.01

Table B5: Effect of Older Sibling Being Born After Cutoff on Younger Sibling's ADHD Outcomes, by Gender

	(1) Boy-Boy	(2) Boy-Girl	(3) Girl-Boy	(4) Girl-Girl
Panel A: ADHD Diagnoses				
Older sibling born after cutoff	-0.0037	-0.0073***	-0.0077*	-0.0050**
[Spillover Effect]	(0.003)	(0.002)	(0.004)	(0.002)
Younger sibling born after cutoff	-0.0171***	-0.0099**	-0.0198***	-0.0147***
[Own Relative Age Effect]	(0.003)	(0.003)	(0.003)	(0.004)
Mean(Y)	0.057	0.038	0.060	0.041
N	71,743	68,388	67,588	63,669
Panel B: ADHD Drugs				
Older sibling born after cutoff	-0.0055**	-0.0065**	-0.0079***	-0.0022
[Spillover Effect]	(0.002)	(0.002)	(0.002)	(0.002)
Younger sibling born after cutoff	-0.0173***	-0.0110***	-0.0214***	-0.0155***
[Own Relative Age Effect]	(0.003)	(0.003)	(0.003)	(0.004)
Mean(Y)	0.064	0.044	0.067	0.048
N	71,743	68,388	67,588	63,669

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more details about the sample, specifications, and control variables. The dependent variable in Panel A is an indicator equal to one if the younger sibling ever has an outpatient claim with an ADHD diagnosis. The dependent variable in Panel B is an indicator equal to one if the younger sibling has at least one ADHD drug claim in the prescription drug data. Column (1) restricts our main sample to sibling pairs in which both the older and younger siblings are male; column (2) uses sibling pairs in which the older sibling is male and the younger sibling is female; column (3) uses sibling pairs in which the older sibling is female and the younger sibling is male; column (4) uses sibling pairs in which the older and younger siblings are female. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B6: Effect of Older Sibling Being Born After Cutoff on Younger Sibling's ADHD Outcomes Measured Before and After High School Graduation

	Diagnoses		Drugs	
	(1)	(2)	(3)	(4)
	Before HS graduation	After HS graduation	Before HS graduation	After HS graduation
Older sibling born after cutoff	-0.0051***	-0.0007	-0.0045***	-0.0011**
[Spillover Effect]	(0.001)	(0.001)	(0.001)	(0.000)
Younger sibling born after cutoff	-0.0112***	-0.0043***	-0.0128***	-0.0036***
[Own Relative Age Effect]	(0.002)	(0.000)	(0.001)	(0.000)
Mean(Y)	0.042	0.007	0.047	0.009
N	271,388	271,388	271,388	271,388

Notes: Each column reports results from a separate regression. See notes under Table 2 for more details about the sample, specifications, and control variables. The dependent variables in columns (1) and (3) are indicators equal to one if the younger sibling's first ADHD diagnosis and ADHD drug claim are observed before graduating from high school, respectively, while the dependent variables in columns (2) and (4) are indicators equal to one if the younger sibling's first ADHD diagnosis and drug claim are observed after high school graduation, respectively. Robust standard errors are clustered on the older sibling's month of birth. Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B7: Effect of Older Sibling Being Born After Cutoff on Younger Sibling's ADHD Diagnoses and Drug Treatment, by Share Funding for Special Needs

	Diagnoses		Drugs		
	(1) Main Sample	(2) Age Diff ≤ 3	(3) Main Sample	(4) Age Diff ≤ 3	(5) Balanced Timing
Older sibling born after cutoff	-0.0016	-0.0049***	-0.0025*	-0.0052**	-0.0016
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Younger sibling born after cutoff	-0.0125***	-0.0101***	-0.0131***	-0.0110***	-0.0120***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Funding × Older sib born after cutoff	-0.0008***	-0.0005	-0.0006**	-0.0003	-0.0007**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Funding × Younger sib born after cutoff	-0.0007**	-0.0012***	-0.0005	-0.0007	-0.0005*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean(Y)	0.049	0.047	0.056	0.053	0.043
N	239,052	154,493	239,052	154,493	239,052

Notes: Each column reports results from a separate regression. See notes under Table 2 for more details about the sample, specifications, control variables, and outcomes. All regressions additionally include interaction terms with the share of school funding for special needs education in the mother's municipality of residence when the younger sibling is aged 7 years old, as well as municipality fixed effects. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B8: Effect of Older Cousin Being Born After Cutoff on Younger Cousin's ADHD Diagnoses and Drug Treatment, Same versus Different Municipalities

	Same Municipality				Different Municipalities			
	Diagnoses	Drugs			Diagnoses	Drugs		
		Main Sample	Age Diff ≤ 3	Bal Tim		Main Sample	Age Diff ≤ 3	Bal Tim
Older cousin born after cutoff	-0.0066***	-0.0084***	-0.0094***	-0.0102***	-0.0059**	-0.0028*	-0.0096***	-0.0021*
[Spillover Effect]	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.001)
Younger cousin born after cutoff	-0.0200***	-0.0192***	-0.0148***	-0.0174***	-0.0184***	-0.0203***	-0.0149***	-0.0215***
[Own Relative Age Effect]	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Mean(Y)	0.047	0.052	0.047	0.040	0.046	0.054	0.051	0.040
N	86,225	86,225	54,022	86,225	116,026	116,026	69,343	116,026

Notes: See the notes under Tables 2 and 3 for more details about the sample, specifications, control variables, and outcomes. Here, we split the sample into cousin pairs that reside in the same municipality and those that reside in different municipalities. Robust standard errors are clustered on the older child's month of birth. Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B9: Effect of Older Sibling Being Born After Cutoff on Younger Sibling's ADHD Diagnosis, Conditional on Older Sibling Having ADHD Diagnosis

	Younger Sibling's ADHD Diagnosis	
	(1)	(2)
	Main Sample	Age Diff ≤ 3
Older sibling born after cutoff	0.0062	-0.0059
[Spillover Effect]	(0.014)	(0.015)
Younger sibling born after cutoff	-0.0378**	-0.0114
[Own Relative Age Effect]	(0.017)	(0.026)
Mean(Y)	0.210	0.203
N	10,655	6,979

Notes: See the notes under Table 2 for more details about the sample, specifications, control variables, and outcomes. In this table, the sample is additionally limited to sibling pairs in which the older sibling has an ADHD diagnosis. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * p < 0.1 ** p < 0.05 *** p < 0.01

Table B10: Robustness of Main Results: Additional Controls and Quadratic Specifications

	Younger siblings		Younger cousins	
	(1) Diagnoses	(2) Drugs	(3) Diagnoses	(4) Drugs
Panel A: Additional Controls				
Older sibling born after cutoff	-0.0038**	-0.0034**	-0.0046***	-0.0035**
[Spillover Effect]	(0.001)	(0.001)	(0.001)	(0.001)
Younger sibling born after cutoff	-0.0156***	-0.0164***		
[Own Relative Age Effect]	(0.002)	(0.001)		
Younger cousin born after cutoff			-0.0187***	-0.0197***
[Own Relative Age Effect]			(0.002)	(0.002)
Mean(Y)	0.049	0.056	0.046	0.053
N	271,388	271,388	208,437	208,437
Panel B: Quadratic Polynomial in Running Variables				
Older sibling born after cutoff	-0.0031***	-0.0065***	-0.0028**	-0.0015
[Spillover Effect]	(0.001)	(0.001)	(0.001)	(0.001)
Younger sibling born after cutoff	-0.0188***	-0.0159***		
[Own Relative Age Effect]	(0.003)	(0.003)		
Younger cousin born after cutoff			-0.0224***	-0.0199***
[Own Relative Age Effect]			(0.004)	(0.004)
Mean(Y)	0.049	0.056	0.046	0.053
N	271,388	271,388	208,437	208,437

Notes: Each column in each panel reports results from a separate regression. The first two columns present results for younger siblings, while the second two columns present results for younger cousins. The outcomes are indicators for any ADHD diagnoses and any ADHD drug claims, as noted in the column headers. All regressions include the same controls as in the main specifications (see notes under Table 2). In Panel A, we add the following additional control variables: indicators for each parent ever being married in the year of the older sibling's birth or the following two years, and each parent's age and age squared in the year of the older sibling's birth. In Panel B, we include quadratic polynomials in the running variables (month of birth relative to January for each sibling). Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B11: Robustness of Main Results: "Doughnut-RD" Dropping December and January Births

	Younger siblings		Younger cousins	
	(1) Diagnoses	(2) Drugs	(3) Diagnoses	(4) Drugs
Older sibling born after cutoff	-0.0081***	-0.0052***	-0.0072***	-0.0073***
[Spillover Effect]	(0.002)	(0.001)	(0.002)	(0.001)
Younger sibling born after cutoff	-0.0159***	-0.0162***		
[Own Relative Age Effect]	(0.002)	(0.001)		
Younger cousin born after cutoff			-0.0182***	-0.0194***
[Own Relative Age Effect]			(0.002)	(0.002)
Mean(Y)	0.049	0.056	0.046	0.053
N	228,737	228,737	175,699	175,699

Notes: Each column reports results from a separate regression. The first two columns present results for younger siblings, while the second two columns present results for younger cousins. The outcomes are indicators for any ADHD diagnoses and any ADHD drug claims, as noted in the column headers. All regressions include the same controls as in the main specifications (see notes under Table 2). In this table, we omit all families with older siblings who are born in December or January of any year. Robust standard errors are clustered on the older sibling's month of birth.
Significance levels: * p < 0.1 ** p < 0.05 *** p < 0.01

Table B12: Robustness of Main Results to Using the Wild Cluster Bootstrap for Inference

	Younger siblings		Younger cousins	
	(1) Diagnoses	(2) Drugs	(3) Diagnoses	(4) Drugs
Older sibling born after cutoff	-0.0059***	-0.0055***	-0.0060***	-0.0048**
[Spillover Effect]	[0.0010]	[0.0040]	[0.0030]	[0.0250]
Younger sibling born after cutoff	-0.0155***	-0.0164***		
[Own Relative Age Effect]	[0.0000]	[0.0000]		
Younger cousin born after cutoff			-0.0187***	-0.0198***
[Own Relative Age Effect]			[0.0000]	[0.0000]
Mean(Y)	0.049	0.056	0.046	0.053
N	271,388	271,388	208,437	208,437

Notes: Each column reports results from a separate regression. The first two columns present results for younger siblings, while the second two columns present results for younger cousins. The outcomes are indicators for any ADHD diagnoses and any ADHD drug claims, as noted in the column headers. All regressions include the same controls as in the main specifications (see notes under Table 2). In this table, $p-$values calculated using the wild bootstrap method are reported in brackets.