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

Petra Persson† Xinyao Qiu‡ Maya Rossin-Slater§

October 5, 2022

Abstract

The health care system uses patient family medical history in many settings, and this practice is widely believed to improve the efficiency of health care allocation. This paper provides a counterpoint by documenting that reliance on hereditary information can amplify the misallocation of low-value care. We study Attention Deficit Hyperactivity Disorder, and show that reliance on family medical history generates a "snowball effect"—the propagation of an original marginal diagnosis to a patient's relatives. This snowball effect raises the private and social costs of low-value care.

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, Barton Willage, 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 and Sarah Bögl at the Research Institute for Industrial Economics 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 Health Policy, 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.[2]

At the same time, an emerging literature examines the value of *marginal* diagnoses for 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).[3] In this paper, we document that the use of family medical history can perpetuate marginal diagnoses across family members, thereby raising caseloads and health care costs. The existence of this "snowball effect" suggests that the potential benefit or cost of a single marginal diagnosis is much larger than it would be if it were limited to only the originally diagnosed patient. In particular, if a marginal diagnosis has low or even negative value for patient health (as documented in several cases in prior work), then the presence of marginal diagnosis spillovers amplifies this cost.

We study this issue in the context of Attention Deficit Hyperactivity Disorder (ADHD), the most commonly diagnosed mental health condition among children, affecting nearly ten

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

[2] For example, Evans et al. (2019) find that screening women whose mothers or sisters have been diagnosed with breast cancer increases the likelihood of early detection of cancer and improves survival rates.

[3] 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., 2020) 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."

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 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).[4]

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.[5] 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.[6] We use population-level Swedish administrative data on children born between July 1, 1985 and June 30, 1996, and start by confirming this previously documented phenomenon in our data and sample with a regression discontinuity (RD) design. We find that children who are born just before the Swedish school entry cutoff of January 1st are 18.6 percent more likely to be diagnosed with ADHD and 17.1 percent more likely to be treated with ADHD medication than their counterparts born just after the cutoff.

Since a child's relative age for grade can influence the outcomes of other family members in the same household through various channels (Landersø et al., 2019; Karbownik and Özek, 2019), we use *non-nuclear* family members to study spillovers of relative-age-induced marginal ADHD diagnoses. Specifically, we focus on first cousins, and show that younger cousins of children born just before the school entry cutoff are 10.0 and 6.5 percent more likely to be diagnosed with ADHD and treated with ADHD drugs, respectively, than the younger cousins of children born shortly after the cutoff. Importantly, these discontinuities exist *conditional*

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

[5] 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.

[6] Age-for-grade is not the only characteristic with respect to which there may be over- or under-diagnosis of ADHD. 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).

on the focal children's own relative age for grade. In fact, the magnitudes of the estimated spillover effects on cousins amount to more than one-fifth of the focal children's own relative age effects on ADHD diagnoses and drug treatment. Additional analysis using information on diagnosis dates provides further support for the spillover mechanism: in cousin pairs in which the older one is born before the cutoff, both cousins are more likely to be diagnosed with the younger one diagnosed *after* the older one; in cousin pairs in which the younger one is born before the cutoff, both are more likely to be diagnosed with the younger one diagnosed *before* the older one.

Next, we examine longer-term educational and labor market outcomes of the younger cousins. Conditional on their own relative age for grade, younger cousins of children who are born before the cutoff—i.e., those who are disproportionately more likely to be diagnosed with and treated for ADHD—have a 0.6 percent lower high school grade point average (GPA), are 0.9 percent less likely to graduate high school on time, and are 2.1 percent less likely to enroll in college by age 21 than younger cousins of children born after the cutoff. We also see a negative spillover coefficient on younger cousins' average earnings at ages 27–30, but it is not statistically significant at conventional levels.[7] These results suggest that the younger cousins of marginally diagnosed children are no better off in the long-term; if anything, there could be long-term human capital costs associated with marginal ADHD diagnosis spillovers.

The existence of the ADHD diagnosis spillovers suggests that information about the older cousin's diagnosis is used in the process through which the younger cousin'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. "Hereditary tagging" in the referral stage would imply that an ADHD diagnosis of one child raises the likelihood that a family member receives a referral for an ADHD screening. While both schools and families play important roles in requesting ADHD screenings in general, we argue that intra-family communication between parents—who are siblings themselves—is a likely key channel in driving spillovers across cousins, who rarely

[7] In all of our models, we control for the younger cousin'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).

attend the same schools.[8] Thus, the ADHD spillovers that we document appear to in part materialize due to a family-driven "referral gap" between the younger cousins of marginally-diagnosed and marginally-undiagnosed children.

In the screening stage, the physician's diagnostic technology plays a critical role. For health conditions for which there is a precise diagnostic technology (e.g., a genetic test), the "referral gap" would not necessarily translate into a "diagnosis gap," as the physician would be able to accurately differentiate between those who do and do not have the condition. 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 relatives of previously diagnosed patients, the "referral gap" will lead to a "diagnosis gap." For ADHD, the screening protocol indicates that physicians should take into account family history of ADHD, but does *not* prescribe differentiating the significance of this information based on the previously diagnosed family member's relative age for grade. Therefore, when physicians follow the protocol, the "referral gap" between the younger cousins of marginally-diagnosed and marginally-undiagnosed children generates spillovers of marginal ADHD diagnoses.

In sum, our two new empirical facts—(i) that marginal ADHD diagnoses propagate across cousins, and (ii) that the value of these spillover diagnoses appears to be negligible (at least in terms of later educational and labor market 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]

[8] We do not have information about which schools children in our data attend, and thus cannot calculate the share of cousins who are in the same or different schools. That being said, the majority of the cousin pairs in our sample live in different municipalities from each other (and therefore definitely attend different schools). The spillover effect is strong even in the sample of cousin pairs who live in different municipalities.

[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 date of birth. In the case of

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., 2020; 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 across 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 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 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. Sibling 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

diabetes, for example, the information could be gleaned from data on the relative's blood sugar level relative to the diagnosis threshold (Alalouf et al., 2019).

[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).

educational outcomes (see, e.g., Fletcher et al., 2012; Breining, 2014; Parman, 2015; Yi et al., 2015; Black et al., 2020; Breining et al., 2019; Daysal et al., 2021), arguing that shifts in parental resource allocation across siblings and within-family infectious disease spread may be important mechanisms. Another related study finds that a Turkish national vaccination campaign targeting children under five years old has spillover effects on vaccine take-up among ineligible older siblings (Alsan, 2017).[11] Our paper provides novel evidence of health-related spillovers across non-nuclear family members—cousins—and also relates to Chen et al. (2022) and Finkelstein et al. (2022), 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

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.

Referral for ADHD screening. In order to be screened for ADHD, an individual needs to see a specialist provider (a psychiatrist). Referrals to screenings can be initiated by both parents and schools.[14]

[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.

[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] All children attending school in Sweden receive free annual health check-ups at school, 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

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 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.

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

right to additional evaluations for any health issues detected in these screenings at school.

[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.

Sweden.

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.[18] With normal progression (i.e., no grade retention), students graduate high school in the year they turn 19 years old. Students can enroll in college after they graduate high school, but it is common for Swedes to take a "gap year." Thus, standard age at college enrollment is around 21 years.[19]

3 Data and Sample

We link several sources of data for our analysis: the universe of Swedish birth records, outpatient, and prescription drug claims data from the National Board of Health and Welfare (NBHW; in Swedish *Socialstyrelsen*), and population register data from Statistics Sweden containing demographic and labor market information. Additionally, we have a data set from Statistics Sweden with the identifiers of all first cousins of each individual. The birth records data cover all births in Sweden from 1985 to 2017; the population register data are available annually from 1990 to 2019 (with the exception of information about the high school GPA, which is available through 2016); the outpatient records are available for years 2001–2016; and prescription drug claims are available for the period July 2005 to December 2017.

The birth records data contain detailed information on pregnancy and birth outcomes, including gestational age in days and expected due date. These data allow us to compute exact dates of birth for the children in our analysis.[20]

We then construct a sample containing pairs of first cousins as follows. We start with the universe of children born in Sweden between 1985 and 2001 with information on exact date of birth. For every child, we find all of his/her first cousins using the Statistics Sweden data

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

[19] According to recent data, only 30 percent of Swedish 19-year-old women and 20 percent of Swedish 19-year-old men applied to college in 2019 (SCB, 2019).

[20] Specifically, we subtract 280 days (40 weeks) from the expected due date to obtain the conception date, and then add the gestational age in days to obtain the actual date of birth. We then compare the resulting month and year of birth to the month and year of birth reported in the population register data, and drop all observations in which the calculated and reported birth-month-years do not match.

set with first cousin identifiers. We then construct pairs of cousins, where each pair has an older and a younger child.[21] We keep pairs with an age difference of no more than 5 years. Finally, we restrict the sample to pairs in which the older cousin is born between July 1985 and June 1996. This final sample restriction allows us to have a sample that is balanced in the running variable of our RD analysis, which is the older cousin's date of birth relative to the school entry cutoff of January 1.[22] Our final analytic sample consists of 1,122,772 cousin pairs, among which there are 575,224 unique older cousins and 616,242 unique younger cousins. As we discuss below, our primary empirical specifications use a 75-day bandwidth around the school entry cutoff, which yields a sample of 432,696 cousin pairs.

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).[23] We also study ADHD drug treatment using the prescription drug data. We create 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.[24]

We additionally use the population register data to study educational and labor market outcomes of the younger cousins. When studying educational outcomes, we limit the sample to pairs in which the younger cousin is born in 1997 or earlier, and consider three measures: cumulative high school GPA (from the 2016 data), an indicator for graduating high school on time (i.e., no later than the year in which a child turns 19 years old), and an indicator for enrolling in college by age 21. To study adult earnings, we limit the sample to pairs in which

[21] Note that a child can be both an older cousin and a younger cousin, and a child can be the younger (older) cousin of multiple different older (younger) cousins. We drop cousin pairs in which both cousins are born in the same year-month.

[22] Our analysis essentially aggregates RDs across 11 fiscal years of birth of the older child, from July 1985–June 1986 through July 1995–June 1996. We include fixed effects for the older child's fiscal year of birth in our regression models.

[23] See: `https://www.icd10data.com/ICD10CM/Codes/F01-F99/F90-F98/F90-/F90`. We do not have ICD codes at a higher level of detail (i.e., more digits) to further separate into various types of ADHD (e.g., inattentive, hyperactive, or combined types).

[24] Specifically, we consider all drug claims with Anatomical Therapeutic Chemical (ATC) codes that start with "N06BA" except "N06BA07", as well the ATC code "C02AC02". Note that the diagnosis and drug indicators are not identical for two reasons. First, some children with a diagnosis do not receive prescription drug treatment. Second, our prescription drug records go through 2017, while the outpatient data end in 2016; thus, we are missing an entire year of potential diagnoses for children whose drug claims we can observe. In addition, there is likely some under-reporting of diagnoses in the outpatient data.

the younger cousin is born in 1992 or earlier, and calculate average annual earnings over ages 27–30.[25]

The population register data also provides us with a number of family-level control variables, including whether each parent of each cousin is foreign-born, parental education level, and household income in each cousin's household.[26]

Sample means. Appendix Table B1 presents sample means of some of the key variables in our analysis, separately for the older and younger cousins of each pair in panels A and B, respectively. The first column uses the entire analysis sample, while the second and third columns are split into pairs with older cousins born in July-December ("Before Cutoff") and January-June ("After Cutoff"), respectively. About 3.6 (4.0) and 4.3 (4.8) percent of older and younger cousins in our sample have an ADHD diagnosis (ever have an ADHD drug claim), respectively. Around six percent of fathers and 5.3 percent of mothers are foreign-born, and approximately 12 and 11 percent of fathers and mothers have college degrees, respectively. The average older cousin has about 2.0 younger cousins, and mean birth spacing between cousin pairs is about 29 months. Cousin pairs with older cousins 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 somewhat higher among children in pairs with July-December than January-June older cousin births.

4 Empirical Design

Our goal is to analyze how a marginal ADHD diagnosis among older children affects their younger cousins' ADHD-related outcomes. To do so, we leverage the discontinuity in the older

[25] For cohorts for whom earnings are not observed at a particular age in this age range (e.g., we do not observe earnings at ages 28–30 for the 1992 cohort since our data only go through 2019), we calculate the average based on the age(s) we do observe.

[26] We measure each parent's education level in the year of their child's birth, and the household income is the average over the year of the child's birth and the following two years. For children born before 1990 (when the population register data begin), we use information from 1990 for parental education level, and average over years 1990–1992 for household income. In addition, while we also observe parental age and marital status, we do not include these control variables in our models 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 at the time of measurement of marital status than parents of children born in December.

cousin'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.

Since we have information on exact dates of birth in our data, we use the older cousin's day of birth relative to January 1 as the running variable in our RD models.[27] Our primary specifications use a bandwidth of 75 days around the cutoff and a linear spline parametrization; in Appendix B, we explore the sensitivity of our estimates to other bandwidths and higher-order polynomials, as well as to non-parametric RD models with optimal bandwidth selection algorithms (Calonico et al., 2014a,b, 2017, 2018).

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

$$ADHD_i = \alpha_0 + \alpha_1 \mathbf{1}[D_i < c] + f(D_i - c) + \mathbf{1}[D_i < c] \times f(D_i - c) + \mathbf{x'}_i\kappa + \epsilon_i \quad (1)$$

for every older cousin i in our analysis sample. $ADHD_i$ is an ADHD-related outcome (i.e., an indicator for either a diagnosis or drug take-up). c denotes January 1st, the school entry cutoff date. The variable $\mathbf{1}[D_i < c]$ is an indicator for the older cousin i being born within the July 1–December 31 window (i.e., *before* the cutoff, and thus relatively young-for-grade), and zero otherwise. $f(D_i - c)$ is a linear function of the running variable, the older cousin's day of birth centered around January 1, which we allow to have different slopes on opposite sides of the cutoff. We show results with and without controls in vector $\mathbf{x}_i$, which includes an indicator for whether the older cousin is male, the total number of cousins in the family, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the older cousin's birth (high school only, some college, college degree or more), the log household income of the family averaged over the year of the older cousin's birth and the following two years, and fixed effects for fiscal years (July–June) of birth of the older cousins.

Then, to study spillover effects on younger cousins' outcomes, we estimate models of the

[27] Following Lee and Card (2008)'s guidance on RD estimation with a discrete running variable, we cluster standard errors on the running variable (i.e., the older child's day of birth).

following form:

$$\begin{aligned} Y_{ij} = \beta_0 + \beta_1 \mathbf{1}[D_i < c] + f(D_i - c) + \mathbf{1}[D_i < c] \times f(D_i - c) \\ + \beta_2 \mathbf{1}[D_j < c] + f(D_j - c) + \mathbf{1}[D_j < c] \times f(D_j - c) \\ + \mathbf{x'}_{ij}\pi + \varepsilon_{ij} \end{aligned} \tag{2}$$

for each pair of older cousin i and a younger cousin j. Y_{ij} is an outcome of interest, such as an indicator for the younger cousin having an ADHD diagnosis. In addition to the variables capturing the older cousin's day of birth relative to January 1st that are the same as in equation (1), we control for the younger cousin's *own* relative age for grade by including analogous variables based on the younger child's day of birth centered around January 1st: $\mathbf{1}[D_j < c]$ and $f(D_j - c)$. As with model (1), we show results with and without controls in $\mathbf{x}_{ij}$, which now includes the following pair-level variables: cousin birth spacing (in months), indicators for whether the older and younger cousin is male, the total number of cousins in the family, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of each child's birth, the log household income of each cousin's household averaged over the first three years of the child's life, and fixed effects for fiscal years of birth of the older and younger cousins. The main coefficient of interest is β_1, which represents the difference in outcomes between younger cousins in pairs in which the older cousins are born before and after January 1 in every fiscal year, holding constant the younger cousin's own birth day 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 day of birth relative to January 1, and the treatment variable is an indicator for the older child being born in the second half of the year, and thus relatively younger-for-grade than his/her counterparts born in the first half of the year. This treatment variable, in turn, generates a discontinuity in ADHD diagnoses among the older cousins. 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 cousin's relative age for grade might influence his/her younger cousins' outcomes.

To assess issue (1), we begin by plotting a histogram of births at a daily level using our sample of older cousins in Appendix Figure B1 and 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 8.11.

To investigate this difference in the density of births between December and January further, 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. Appendix Figure B2(a) plots the average length of gestation in days by the child's birth week among our sample of older cousins, while Appendix Figure B2(b) plots the younger cousins' average gestation by their older cousin's week of birth. There are no visible discontinuities in mean gestation length at the January 1 cutoff in either graph. Column (1) of Appendix Table B2 reports results from estimating models (1) and (2) using the older and younger cousins' gestation length in days as the outcome, respectively, with no indication of any significant discontinuity at the school entry cutoff.

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 study whether there are any differences in family background characteristics between cousin pairs in which the older cousins are born before and after the cutoff. We do not observe any discontinuous changes in average birth spacing between cousins, the gender composition of older and younger cousins, or parental education levels in Appendix Figures B3, B4, and B5, respectively. Appendix Figure B6 plots *predicted* ADHD diagnosis and drug treatment indicators of the older and younger cousins by the birth week of the older cousin. The predicted variables are constructed by regressing each of the ADHD outcomes on the control variables included in $\mathbf{x}_{ij}$ in equation (2) except for the fiscal year

fixed effects.[28] We do not see any discontinuities in these predicted outcomes at the school entry cutoff. Lastly, Columns (2)–(5) of Appendix Table B2 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 cousin being born before versus after the January 1 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 B1 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 cousin pairs with older cousins born in a two-week bandwidth surrounding the cutoff, which yields similar results to those from our baseline models (see Appendix Table B4).

When it comes to issue (2), we recognize that the spillover effects of an older child's relative age for grade on his/her younger relative'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 than their counterparts with older siblings who are youngest for their grade.[29] However, these family dynamics are less likely to be relevant for non-nuclear family members who do *not* share a household or the same set of parents. We therefore believe that our focus on 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.

[28] Results are similar if we include fiscal year fixed effects in the prediction models.

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

5 Results

We begin by using our analysis sample of older cousins 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 cousins' relative-age-induced marginal diagnoses of ADHD on their younger cousins' likelihoods of ADHD diagnosis and drug treatment. To assess whether there are any human capital-related costs or benefits associated with these spillovers, we also study younger cousins' long-term educational and labor market outcomes. Lastly, we discuss possible mechanisms driving these marginal ADHD diagnosis spillovers.

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

Figure 2 plots ADHD-related outcomes of the older cousins by their *own* birth week (centered around the week that begins with January 1). 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 at the cutoff—children who are youngest for their grade (i.e., born shortly before January 1) 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 column (1) is an indicator equal to one if the older cousin ever has an outpatient claim with an ADHD diagnosis, while the outcome in column (2) is an indicator equal to one if the older cousin has at least one ADHD drug claim. We present results from models with and without control variables in Panels A and B, respectively. Across both columns and panels, we find that being born before the school entry cutoff is associated with a significantly higher likelihood of being diagnosed with and treated for ADHD. Specifically, focusing on the

estimates from Panel A, we observe that children born before the cutoff are 0.7 percentage points more likely to both be diagnosed with ADHD and treated with ADHD drugs. Relative to the corresponding sample means, these estimates yield effect size magnitudes of 18.6 and 17.1 percent, respectively.

Panel A of Appendix Table B3 shows the sensitivity of our estimates to using different polynomials in the running variable, while sub-figures (a) and (b) of Appendix Figure B7 show results using different bandwidths. In addition, Panel A of Appendix Table B4 presents results from "doughnut-RD" models, which omit cousins born in a two-week bandwidth surrounding the cutoff to address the issue of potentially non-random sorting around the threshold. The discontinuity in the likelihood of ADHD diagnosis and ADHD drug treatment between children born just before and just after the school entry cutoff is robust across these specification choices.

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.[30] 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 children born shortly before the January 1 cutoff relative to those born shortly after the cutoff? 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

[30]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.

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 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.

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 B8, 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 their younger cousins.[31]

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

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

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 cousin. Not surprisingly, the younger cousin's *own* relative age for grade has an effect on the probability of ADHD diagnosis and drug treatment—columns (1) and (2) of Panel A show that younger cousins born before the school entry cutoff are 1.4 and 1.5 percentage points more likely to have an ADHD diagnosis and an ADHD drug claim, respectively, than those born after the cutoff. However, conditional on the younger cousin's own relative age for grade, we also see a significant effect of the older cousin's relative age for grade on the younger cousin's likelihood of ADHD diagnosis and drug

[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.

treatment. Younger cousins of older children who are born before the school entry cutoff are 0.4 percentage points more likely to have an ADHD diagnosis and 0.3 percentage points more likely to have an ADHD drug claim, corresponding to 10.0 and 6.5 percent effect sizes when evaluated at the respective dependent variable means. The magnitudes of the spillovers on ADHD diagnoses and drugs are 30.9 and 20.1 percent, respectively, of the sizes of the younger cousin's own relative age-for-grade effects on these outcomes.[32]

Panel B of Appendix Table B3 shows the sensitivity of our spillover estimates to using different polynomials in the running variables, while sub-figures (c) and (d) of Appendix Figure B7 present the spillover results based on RD models with different bandwidths. Additionally, Panel B of Appendix Table B4 reports estimates from "doughnut-RD" models, which omit cousin pairs in which the older cousins are born in a two-week bandwidth surrounding the cutoff. Finally, Appendix Table B5 shows results from RD models with local linear polynomials that use different optimal bandwidth algorithms to select the bandwidths of the number of days used on each side of the school entry cutoff.[33] While the magnitudes of the spillover effects vary somewhat across the specifications, they are actually strongest in models that use narrower bandwidths (in columns 6 to 10).

In Table 3, we take advantage of information about the dates of ADHD diagnoses in the outpatient data to shed light on the timing patterns that might be consistent with potential spillover effects. We estimate the same models as in Table 2, studying three additional outcomes across the three columns of the table: (1) an indicator equal to 1 if both the older and younger cousin are diagnosed at some point in our data, (2) an indicator equal to 1 if both cousins are diagnosed, and the younger cousin is diagnosed after the older cousin, and (3) an indicator equal to 1 if both cousins are diagnosed, and the older cousin is diagnosed

[32] We have also explored heterogeneity in spillover effects across cousin pairs with differing gender composition and with respect to various measures of the families' socioeconomic status, finding no consistent patterns.

[33] We use triangular kernels and robust bias-corrected inference procedures in all models. The optimal bandwidth algorithms are: (1) one common mean squared error (MSE)-optimal bandwidth selector for both sides of the cutoff; (2) two different MSE-optimal bandwidth selectors (below and above the cutoff); (3) one common MSE-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof); (4) minimum of (1) and (3); (5) median of (1), (2), and (3) for each side of the cutoff separately; (6) one common coverage error rate (CER)-optimal bandwidth selector; (7) two different CER-optimal bandwidth selectors (below and above the cutoff); (8) one common CER-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof); (9) minimum of (6) and (8); (10) median of (6), (7), and (8) for each side of the cutoff separately. We use the Stata "rdrobust" command for these analyses (Calonico et al., 2017). We report the number of days used in the left and right-hand bandwidths in each model at the bottom of the table.

after the younger cousin. Column (1) shows that both cousins' relative ages-for-grade are predictive of the likelihood that they are both diagnosed. The next two columns indicate a clear timing pattern: the cousin with a birthday shortly before the cutoff appears to be the first to be diagnosed, and is then later followed by the other cousin. Specifically, Column (2) shows that if the older cousin is born before the cutoff, then the likelihood that both cousins are diagnosed *and* the younger cousin is the second to be diagnosed is 0.5 percentage points higher relative to a case where the older cousin is born after the cutoff. Conversely, in column (3) we see that if the younger cousin is born before the cutoff, then the likelihood that both cousins are diagnosed *and* the older cousin is the second to be diagnosed is 0.6 percentage points higher than for a pair with the younger cousin born after the cutoff. These results further support the idea that a relative-age-for-grade-induced marginal ADHD diagnosis can "snowball" through the family tree.

5.3 Impacts on Younger Cousins' Educational and Economic Outcomes

How do the spillovers of marginal ADHD diagnoses influence longer-term educational and labor market outcomes among the younger cousins? In Figure 5 and Table 4, we analyze our three educational outcomes and annual earnings averaged over ages 27–30.

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 cousins who are born before the cutoff have a lower GPA, are less likely to graduate high school on time, are less likely to enroll in college by age 21, and have lower age 27–30 earnings than their counterparts who are born after the cutoff.

When analyzing spillovers, we find that, conditional on own relative age for grade, younger cousins of children who are born before the cutoff have a 7.6 percentage point (0.6 percent) lower high school GPA, are 0.7 percentage points (0.9 percent) less likely to graduate high school on time, and are 0.6 percentage points (2.1 percent) less likely to be enrolled in college than the younger cousins of those who are born after the cutoff. The statistical significance of

these estimates varies slightly depending on whether the full set of control variables is included, but the patterns are similar across the two panels of the table. We also observe a negative spillover effect on earnings, but it is not statistically significant in either model. 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 educational or economic benefits associated with marginal ADHD diagnosis spillovers.

Moreover, our results on these outcomes are complementary to the existing clinical evidence on the impacts of ADHD treatment. A positive ADHD diagnosis 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.[34] To date, we know very little about the longer-term impacts of ADHD drug treatment on measures of individuals' well-being. Our estimates provide suggestive evidence that there could be important long-term human capital costs associated with the use of ADHD drugs on the margin.

5.4 Potential Mechanisms Behind Marginal ADHD Spillovers

We have established that there are substantial spillovers of marginal ADHD diagnoses across cousins, and that there may be long-term human capital costs associated with these spillovers. Now, we discuss the underlying mechanisms that may be generating these spillovers, and 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 cousin affects the likelihood that a younger cousin is diagnosed.

[34] 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).

Referral stage. If "hereditary tagging" takes place in the referral stage, then a child with a family history of ADHD is more likely to be referred to an ADHD screening than a child without such a family history. In our setting, it is possible that the families of younger cousins of marginally diagnosed older children are more likely to request ADHD screenings than the families of younger cousins of marginally undiagnosed older children. Since most cousins do not attend the same schools, it seems that intra-family communication is central to generating spillovers during the referral stage.[35] Indeed, there are multiple reasons for why an older cousin's diagnosis reduces the costs of the family requesting an ADHD screening for the younger cousin. For example, families in which one child is being treated for ADHD have an established connection with a child psychiatrist. In addition, treatment of a one child may make the families less worried about potential adverse consequences of treatment (e.g., side effects of drugs), or make salient the potential advantages of an ADHD diagnosis (e.g., receiving extra time on tests at school).

Diagnosis stage. As a consequence of "hereditary tagging" at the referral stage, more children with an older cousin born just before the school entry cutoff end up in the doctor's office for an ADHD screening than children with an older cousin born right after the cutoff. At this point, the physician's diagnostic technology plays an important role. 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 cousins of pre-cutoff-born children may be over-referred to ADHD screenings compared to the younger cousins of post-cutoff-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 relatives of previously diagnosed patients,

[35] In fact, we find that the ADHD spillovers exist even among cousin pairs who live in different municipalities and therefore certainly do not attend the same schools.

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.

As discussed in Section 2, while the ADHD screening 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. Thus, if physicians follow the protocol, then they do not "undo" the referral gap. Instead, the "referral gap" translates into a "diagnosis gap."

However, as highlighted in Figure 3, the older child's date 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 suggests that there exists a simple adjustment to physician protocol that can mitigate marginal ADHD spillovers: that physicians 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.[36] More broadly, our findings suggest that designing physician protocols that take into account the severity of the condition in the older child —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."

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 the

[36]In fact, if physicians were to incorporate the relationship between relative age for grade and ADHD diagnoses for *all* children (and not just those who have been flagged through hereditary tagging), then the initial relationship between relative age for grade and ADHD would be eliminated.

utilization of (privately or publicly funded) health care 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 commonly dianosed 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 18.6 and 17.1 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 cousins. We find that younger cousins of children born before the cutoff are 10.0 and 6.5 percent more likely to be diagnosed with and treated for ADHD, respectively, than the younger cousins of children born after the cutoff.

To investigate the long-term implications of these diagnosis spillovers, we also show that younger cousins of children born before the school entry cutoff have worse long-run human capital outcomes than the younger cousins of children born after the cutoff. While we cannot completely rule out that other changes in (non-nuclear) 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 cousins.

We argue that intra-family communication is likely to play an important role in the "tagging process" at the referral stage, in which younger cousins 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 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," *The Economic Journal*, 08 2020. ueaa074.

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.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust data-driven inference in the regression-discontinuity design," *The Stata Journal*, 2014, *14* (4), 909–946.

_ , _ , and _ , "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , _ , Max H Farrell, and Rocio Titiunik, "rdrobust: Software for regression-discontinuity designs," *The Stata Journal*, 2017, *17* (2), 372–404.

_ , _ , _ , and _ , "Regression discontinuity designs using covariates," *Review of Economics and Statistics*, 2018, (0).

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," *American Economic Journal: Applied Economics*, 2022.

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.

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. Løken, 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.

Daysal, N. Meltem, Hui Ding, Maya Rossin-Slater, and Hannes Schwandt, "Germs in the Family: The Long-Term Consequences of Intra-Household Endemic Respiratory Disease Spread," Working Paper 29524, National Bureau of Economic Research November 2021.

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 Ostriker, and Heidi Williams, "Screening and selection: The case of mammograms," *American Economic Review*, 2020, *110* (12), 3836–70.

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

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

Evans, D Gareth, S Thomas, J Caunt, A Burch, AR Brentnall, L Roberts, Anthony Howell, M Wilson, R Fox, S Hillier et al., "Final results of the prospective FH02 mammographic surveillance study of women aged 35–39 at increased familial risk of breast cancer," *EClinicalMedicine*, 2019, *7*, 39–46.

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.

Finkelstein, Amy, Petra Persson, Maria Polyakova, and Jesse Shapiro, "A Taste of Their Own Medicine: Guideline Adherence and Access to Expertise =," *American Economic Review: Insights*, 2022.

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.

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.

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.

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.

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.

SCB, "Statistiska Meddelanden Universitet och högskolor," *Sveriges Officiella Statistik*, 2019, (UF 46 SM 1901).

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.

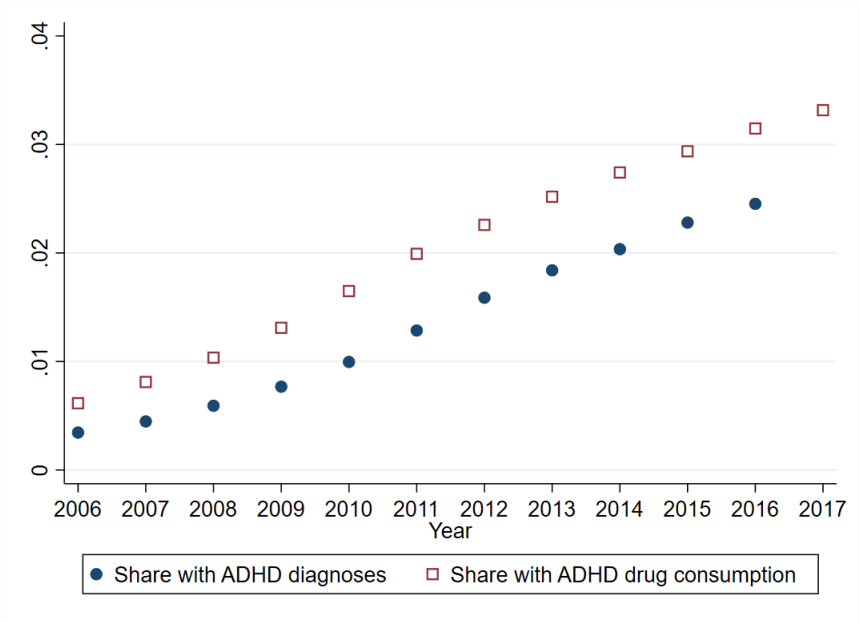
Wilens, Timothy E, Stephen V Faraone, Joseph Biederman, and Samantha Gunawardene, "Does stimulant therapy of attention-deficit/hyperactivity disorder beget later substance abuse? A meta-analytic review of the literature," *Pediatrics*, 2003, *111* (1), 179–185.

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

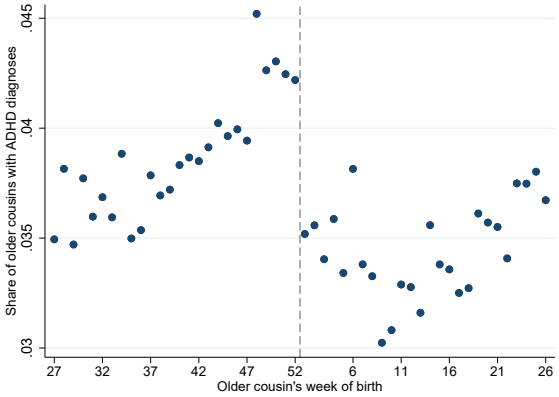
Figure 1: Trends in the Share of Children Ages 6–19 with ADHD



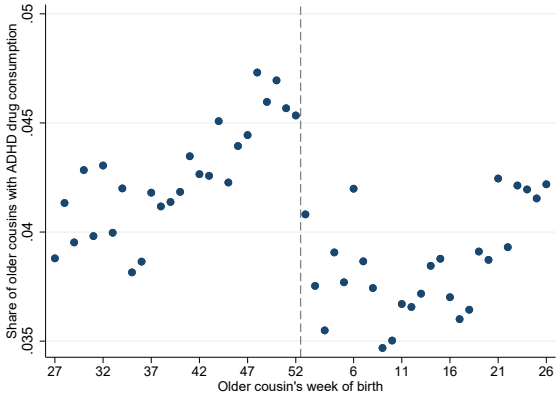
Note: The sample includes children between the ages of 6 and 19 who are born in Sweden, Norway, or the EU 28. For every year, the figure plots the share of these children with at least one ADHD diagnosis in the outpatient data (in blue-filled dots) and at least one ADHD drug claim (in red-outlined squares),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

Figure 2: ADHD Diagnoses and Drug Treatment by Own Week of Birth, Older Cousins Only

(a) Diagnoses

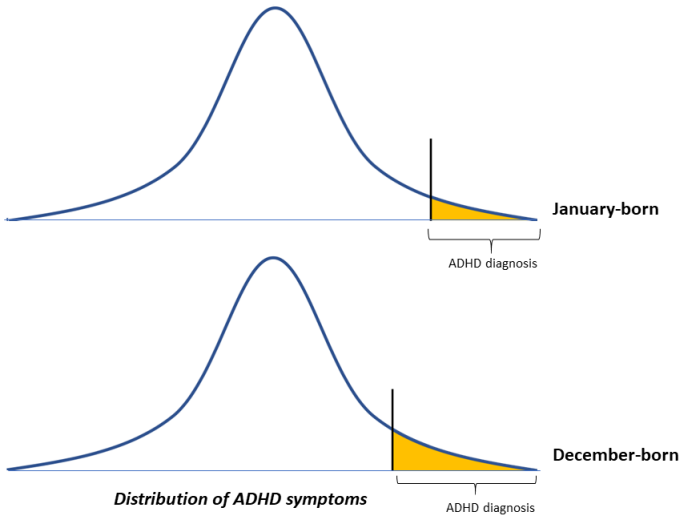


(b) Drug Treatment



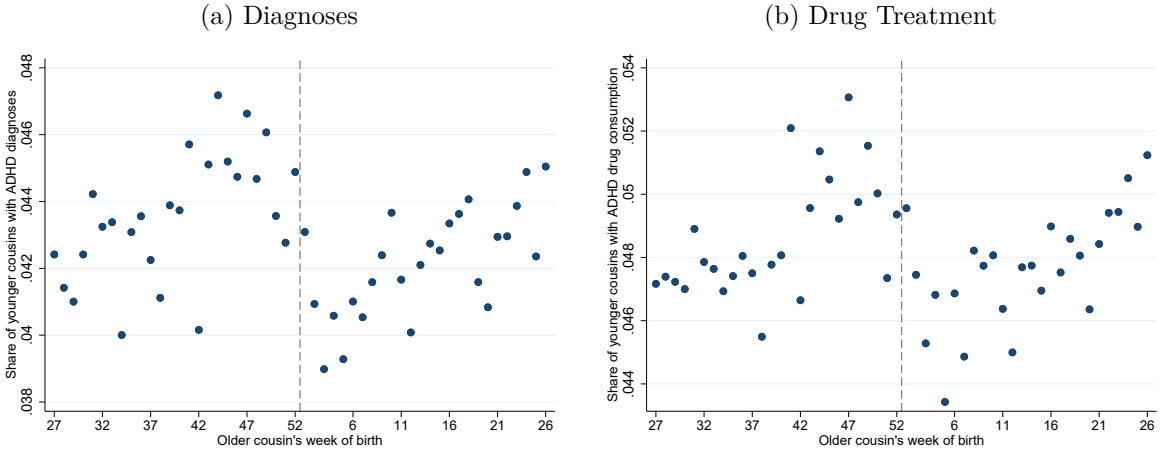
Note: The sample of analysis is the universe of cousin pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. These graphs plot ADHD-related outcomes for older cousins by their own birth week. Sub-figure (a) plots the share of older cousins with an ADHD diagnosis in the outpatient data, while sub-figure (b) plots the share of older cousins 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.

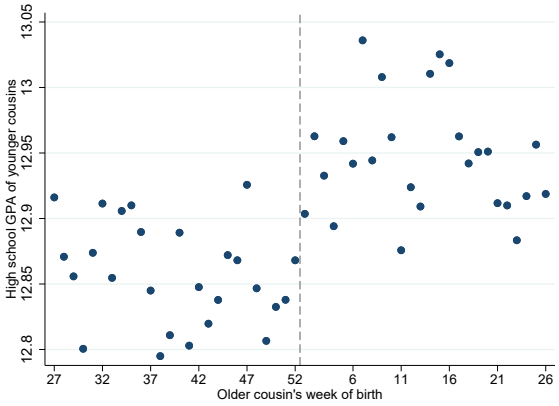
Figure 4: Younger Cousins' ADHD Diagnoses and Drug Treatment by Older Cousin's Week of Birth



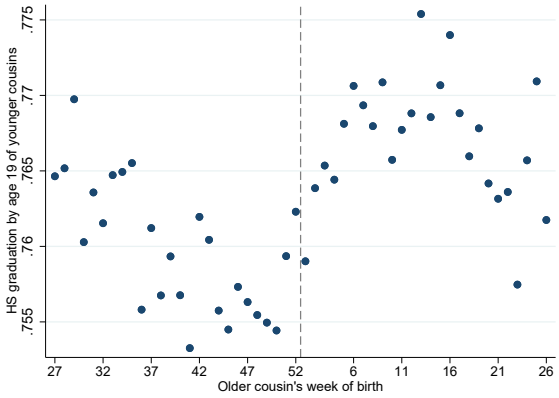
Note: The sample of analysis is the universe of cousin pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. These graphs plot ADHD-related outcomes for younger cousins (on the y-axes) by the birth week of the older cousin (on the x-axes). Sub-figure (a) plots the share 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 5: Younger Cousins' Long-Run Educational and Labor Market Outcomes by Older Cousin's Week of Birth

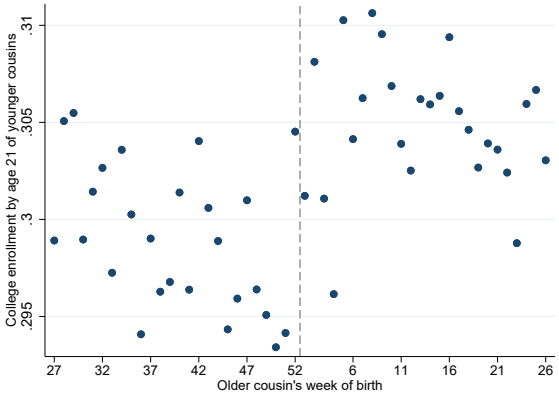
(a) HS GPA



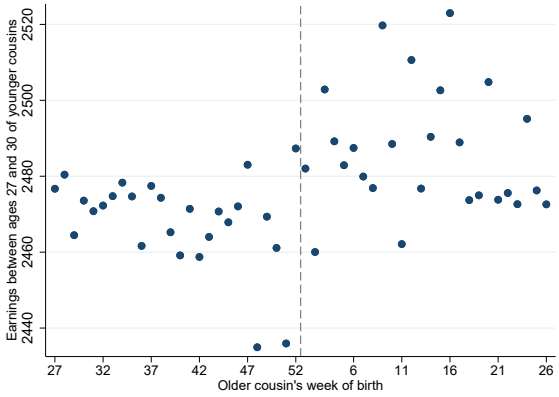
(b) HS Graduation on Time



(c) College Enrollment, by Age 21



(d) Earnings, between Ages 27 and 30



Note: These figures plot average long-run educational and labor market outcomes of younger cousins (on the y-axes), by the birth week of the older cousin (on the x-axes). The sample in sub-figures (a)-(c) is limited to cousin pairs with younger cousins born in 1985-1997 only, and the sample in sub-figure (d) is limited to cousin pairs with younger cousins born in 1985-1992 (see notes under Figure 2 for further description of the main cousins sample). High school GPA is measured in 2016. Graduating high school on time is an indicator set to 1 if an individual graduates from high school no later than the year in which he/she turns 19. College enrollment is an indicator set to 1 if an individual is ever enrolled in college by age 21. Earnings are the work income averaged between age 27 and age 30.

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

	(1) ADHD Diag	(2) ADHD Drug
Panel A: No Covariates		
OC born before the cutoff	0.0069***	0.0070***
[Own Relative Age Effect]	(0.0020)	(0.0019)
Mean(Y)	0.037	0.041
N	221,542	221,542
Panel B: Full Covariates		
OC born before the cutoff	0.0063***	0.0063***
[Own Relative Age Effect]	(0.0019)	(0.0018)
Mean(Y)	0.037	0.041
N	221,542	221,542

Notes: Each column reports results from a separate regression estimating model (1). The sample of analysis is the universe of older cousins born between July 1985 and June 1996, among cousin pairs born in Sweden. The dependent variable in column (1) is an indicator equal to one if the older cousin ever has an outpatient claim with an ADHD diagnosis. The dependent variable in column (2) is an indicator equal to one if the older cousin has at least one ADHD drug claim in the prescription drug data. All regressions have a bandwidth of 75 days and control for a linear spline function for the older cousin's day of birth centered around January 1st (i.e., the running variable in the RD specification). In Panel B, the regressions also include controls for an indicator for whether the older cousin is male, total number of cousins in the family, indicators for whether each parent is foreign-born, indicators for each parent's education categories in the year of the child's birth (high school only, some college, college degree or more), the log household income averaged over the year of the child's birth and the following two years, and fixed effects for the fiscal years of birth of the older cousins. Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

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

	(1) ADHD Diag	(2) ADHD Drug
Panel A: No Covariates		
OC born before the cutoff	0.0043***	0.0031**
[Spillover Effect]	(0.0013)	(0.0014)
YC born before the cutoff	0.0139***	0.0154***
[Own Relative Age Effect]	(0.0013)	(0.0014)
Mean(Y)	0.043	0.048
N	432,693	432,693
Panel B: Full Covariates		
OC born before the cutoff	0.0040***	0.0029**
[Spillover Effect]	(0.0012)	(0.0013)
YC born before the cutoff	0.0117***	0.0127***
[Own Relative Age Effect]	(0.0013)	(0.0014)
Mean(Y)	0.043	0.048
N	432,693	432,693

Notes: Each column reports results from a separate regression estimating model (2). The sample of analysis is the universe of cousins pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. The dependent variable in column (1) is an indicator equal to one if the younger cousin ever has an outpatient claim with an ADHD diagnosis. The dependent variable in column (2) is an indicator equal to one if the younger cousin has at least one ADHD drug claim in the prescription drug data. All regressions control for linear spline functions of the older and younger cousin's day of birth centered around January 1st (i.e., the running variables in the RD specifications). In Panel B, the regressions also include controls for birth spacing (in months), indicators for whether the older and younger cousin is male, total number of cousins in the family, indicators for whether each parent of the older and younger cousin is foreign-born, indicators for each parent's education categories in the year of the child's birth (high school only, some college, college degree or more), the log household income of the older and younger cousin averaged over the year of the child's birth and the following two years, and fixed effects for the fiscal years of birth of the older and younger cousins. Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 3: Timing of ADHD Diagnoses, Older and Younger Cousins

	(1) Both Diag	(2) YC Second	(3) OC Second
Panel A: No Covariates			
OC born before the cutoff	0.0007**	0.0005**	0.0002
[Spillover Effect]	(0.0003)	(0.0002)	(0.0002)
YC born before the cutoff	0.0008**	0.0003	0.0006**
[Own Relative Age Effect]	(0.0004)	(0.0002)	(0.0002)
Mean(Y)	0.003	0.001	0.001
N	432,693	432,693	432,693
Panel B: Full Covariates			
OC born before the cutoff	0.0006*	0.0004*	0.0002
[Spillover Effect]	(0.0003)	(0.0002)	(0.0002)
YC born before the cutoff	0.0006*	0.0001	0.0005**
[Own Relative Age Effect]	(0.0004)	(0.0002)	(0.0003)
Mean(Y)	0.003	0.001	0.001
N	432,693	432,693	432,693

Notes: Each column reports results from a separate regression. The sample and regression specifications are the same as in Table 2. The outcomes are: (1) an indicator equal to 1 if both cousins are diagnosed with ADHD and zero otherwise, (2) an indicator equal to 1 if both cousins are diagnosed and the younger cousin is diagnosed after the older cousin, and (3) an indicator equal to 1 if both cousins are diagnosed and the older cousin is diagnosed after the younger cousin. All regressions control for linear spline functions of the older and younger cousin's day of birth centered around January 1st (i.e., the running variables in the RD specifications). In Panel B, the regressions also include controls for birth spacing (in months), indicators for whether the older and younger cousin is male, total number of cousins in the family, indicators for whether each parent of the older and younger cousin is foreign-born, indicators for each parent's education categories in the year of the child's birth (high school only, some college, college degree or more), the log household income of the older and younger cousin averaged over the year of the child's birth and the following two years, and fixed effects for the fiscal years of birth of the older and younger cousins. Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table 4: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's Long-Run Educational and Labor Market Outcomes

	(1) HS GPA	(2) HS Grad by 19	(3) Enroll by 21	(4) Earning 27-30
Panel A: No Covariates				
OC born before the cutoff	-0.0763*	-0.0065**	-0.0062*	-21.5148
[Spillover Effect]	(0.0395)	(0.0028)	(0.0035)	(15.8552)
YC born before the cutoff	-0.4641***	-0.0723***	-0.0269***	-132.1388***
[Own Relative Age Effect]	(0.0393)	(0.0029)	(0.0030)	(13.4326)
Mean(Y)	12.908	0.763	0.302	2476.763
N	346,634	402,055	401,066	217,469
Panel B: Full Covariates				
OC born before the cutoff	-0.0557	-0.0049*	-0.0035	-19.5195
[Spillover Effect]	(0.0339)	(0.0028)	(0.0029)	(15.7612)
YC born before the cutoff	-0.5073***	-0.0693***	-0.0237***	-124.1077***
[Own Relative Age Effect]	(0.0356)	(0.0029)	(0.0028)	(12.8154)
Mean(Y)	12.908	0.763	0.302	2476.763
N	346,634	402,055	401,066	217,469

Notes: Each column reports results from a separate regression estimating model (2). The sample of analysis is the universe of cousins pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. The sample in columns (1)-(3) is limited to cousin pairs with younger cousins born in 1985-1997, and the sample in column (4) is limited to cousin pairs with younger cousins born in 1985-1992. High school GPA is measured in 2016. Graduating high school on time is an indicator set to 1 if an individual graduates from high school no later than the year in which he/she turns 19. College enrollment is an indicator set to 1 if an individual is ever enrolled in college by age 21. Earnings are the work income averaged between age 27 and age 30. See notes under Table 2 for more details about the sample, specifications, and control variables. Robust standard errors are clustered on the older cousin's day 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.[37] 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.

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

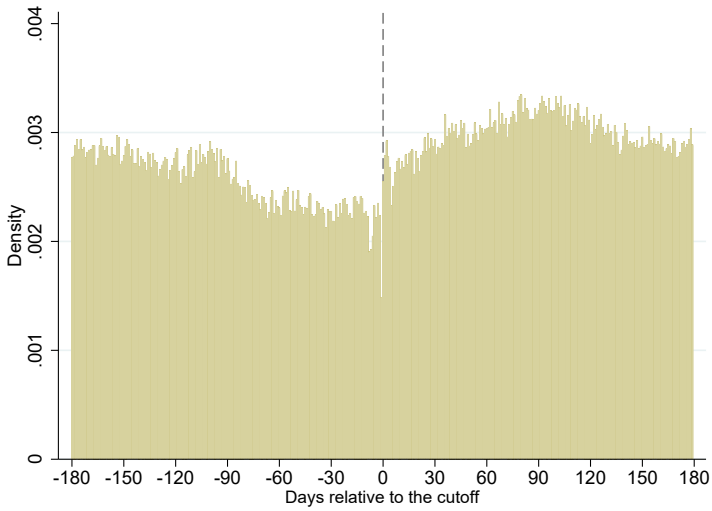
3. Often runs about or climbs in situations where it is not appropriate (adolescents or adults may be limited to feeling restless).
4. Often unable to play or take part in leisure activities quietly.
5. Is often "on the go" acting as if "driven by a motor".
6. Often talks excessively.
7. Often blurts 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.

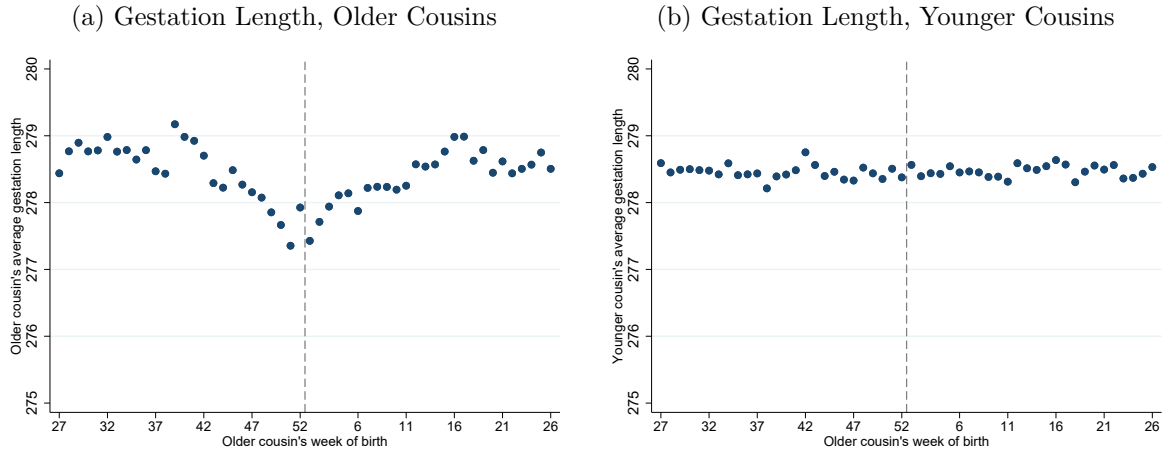
B Additional Results

Figure B1: Distribution of Births at the Daily Level, Older Cousins



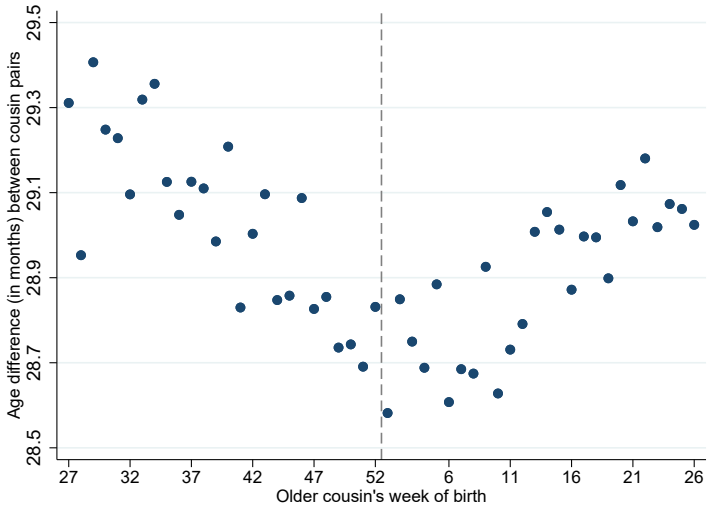
Note: See notes under Figure 2 for more information about the sample. The figure shows a histogram of the distribution of older cousins' births at the daily level, with a bandwidth of 180 days around the cutoff (January 1).

Figure B2: Average Gestation Length by Older Cousin's Week of Birth



Note: See notes under Figure 2 for more information about the sample. Sub-figure (a) plots the older cousins' average length of gestation in days by their own week of birth. Sub-figure (b) plots the younger cousins' average length of gestation in days by older cousin's week of birth.

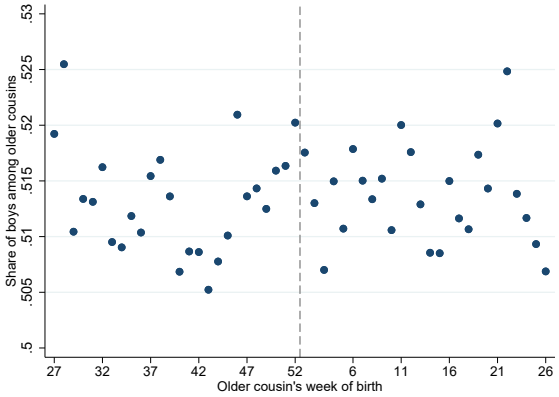
Figure B3: Age Difference Between Cousins by Week of Birth of Older Cousin



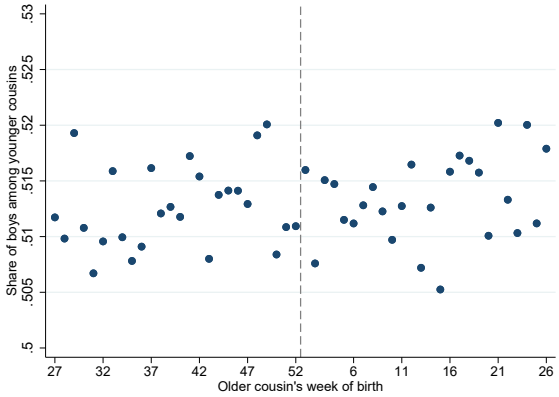
Note: See notes under Figure 2 for more information about the sample. The figure plots the average age difference between cousins (in months) by the birth week of the older cousin.

Figure B4: Cousin Gender Composition by Week of Birth of Older Cousin

(a) Share Boys, Older Cousins



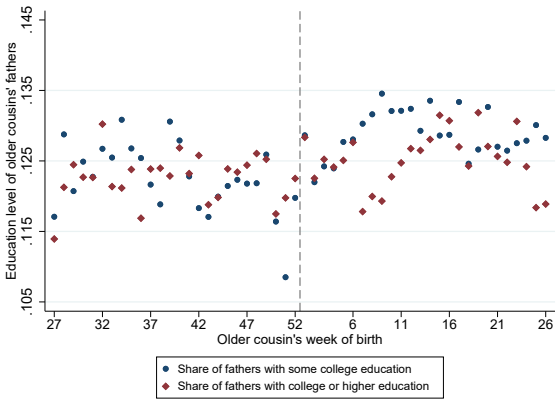
(b) Share Boys, Younger Cousins



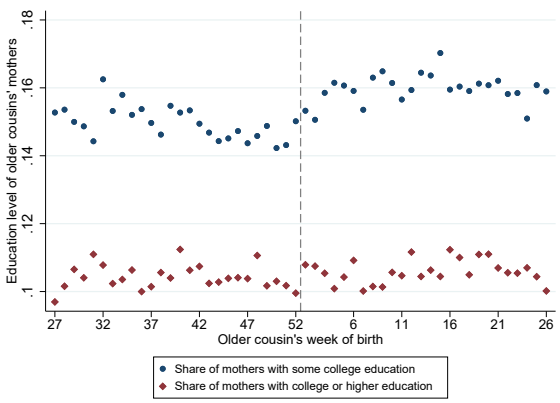
Note: See notes under Figure 2 for more information about the sample. These figures plot the share of boys among older and younger cousins by the birth week of the older cousin. Sub-figure (a) plots the share of boys among older cousins in our main sample of cousin pairs. Sub-figure (b) plots the share of boys among younger cousins in our main sample of cousin pairs.

Figure B5: Older and Younger Cousins' Parental Education Level by Own Week of Birth

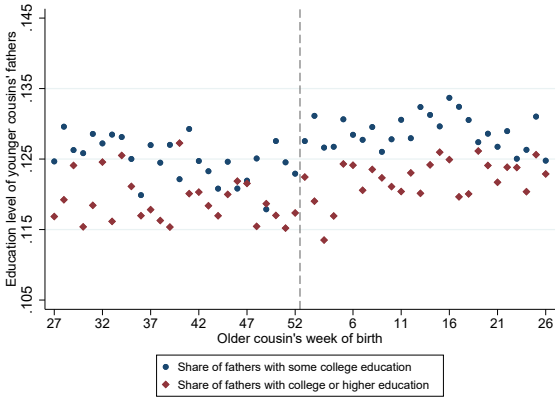
(a) Fathers, Older Cousins



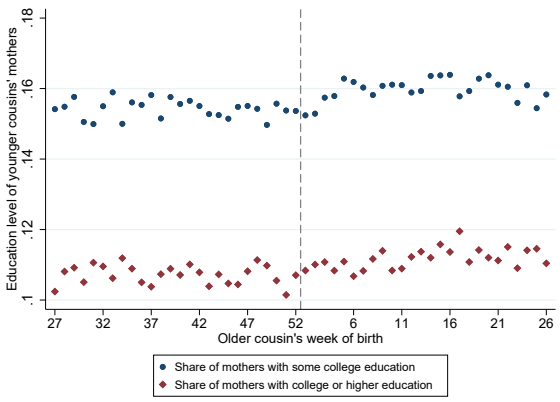
(b) Mothers, Older Cousins



(c) Fathers, Younger Cousins



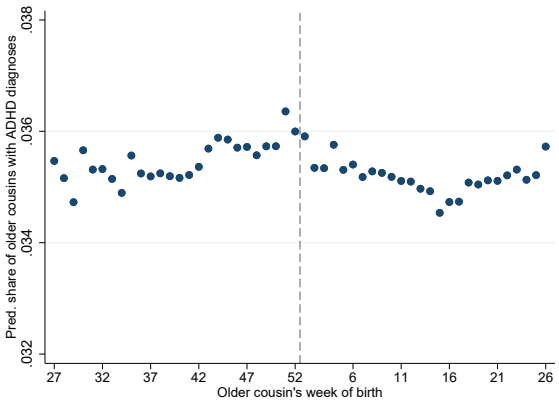
(d) Mothers, Younger Cousins



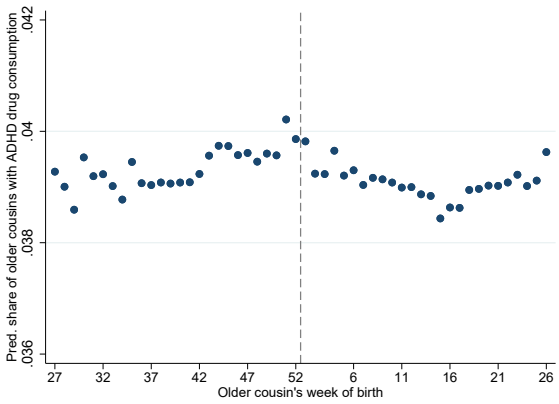
Note: See notes under Figure 2 for more information about the sample. Sub-figure (a) plots the share of older cousins' fathers with some college education (in blue) and the share of older cousins' fathers with college or higher level of education (in red), by the older cousin's week of birth. Sub-figure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

Figure B6: Predicted ADHD Outcomes by Week of Birth of Older Cousin

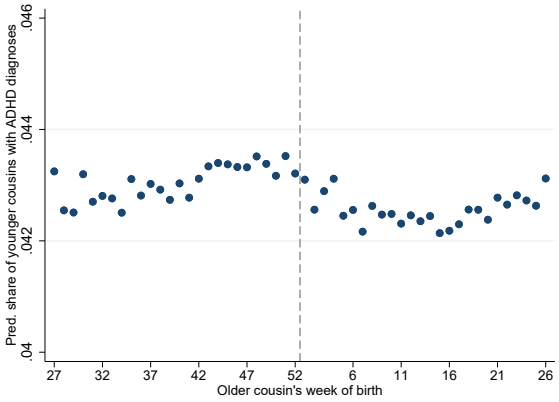
(a) Pred. Diagnoses, Older Cousins



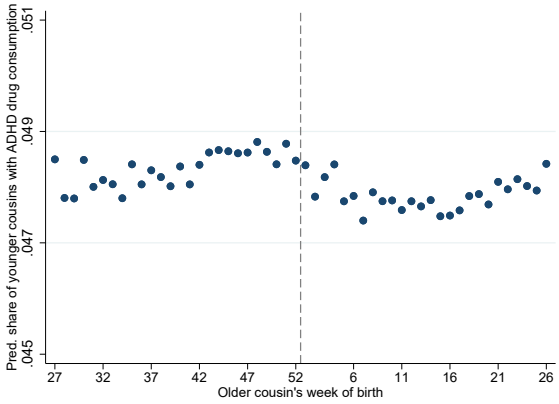
(b) Pred. Drug Treatment, Older Cousins



(c) Pred. Diagnoses, Younger Cousins



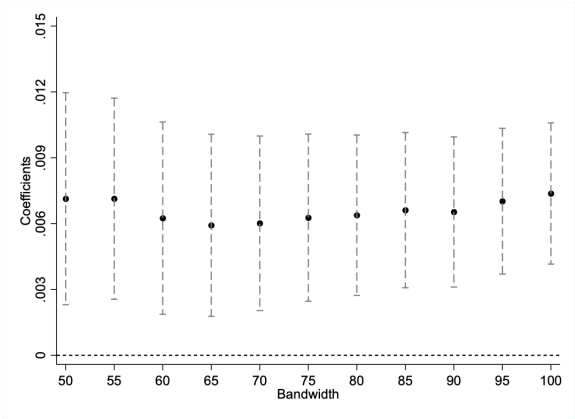
(d) Pred. Drug Treatment, Younger Cousins



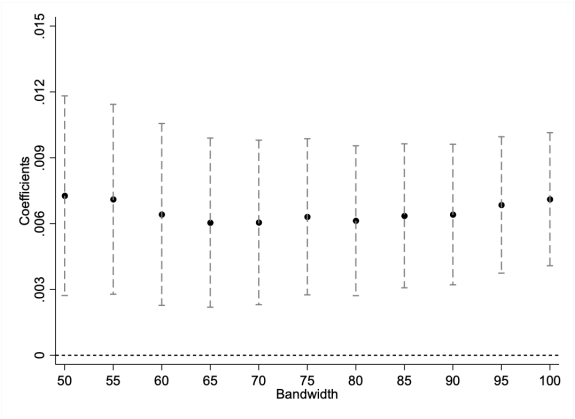
Note: See notes under Figure 2 for more information about the sample. These graphs plot predicted ADHD-related outcomes for the older cousins and younger cousins, respectively, by the birth week of the older cousin. The predicted outcomes of older and younger cousins are constructed by regressing each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

Figure B7: Effect of Older Cousin Being Born Before Cutoff on Own and Younger Cousin's ADHD Diagnoses and Drug Treatment, with Varying Bandwidth

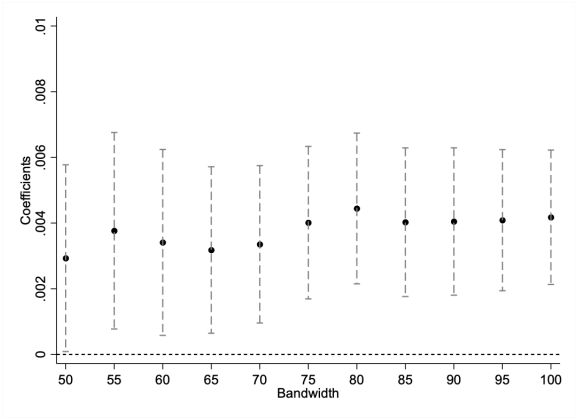
(a) Diagnoses, Older Cousins



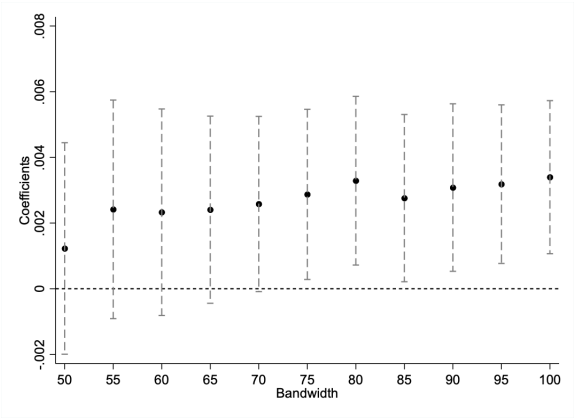
(b) Drug Treatment, Older Cousins



(c) Diagnoses, Younger Cousins



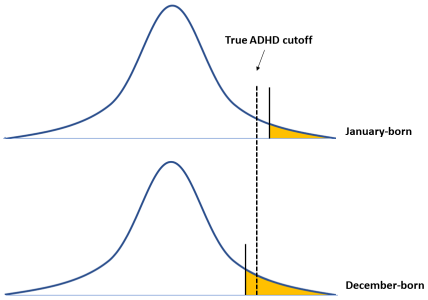
(d) Drug Treatment, Younger Cousins



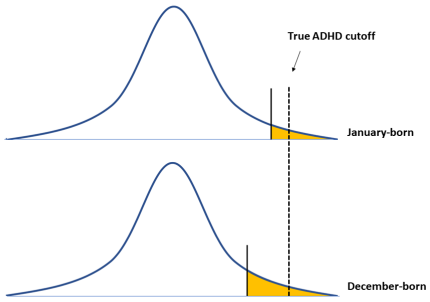
Note: These figures plot regression coefficients and 95% confidence intervals when we include the same full set of control variables but vary the bandwidth from 50 days to 100 days. See notes under Table 2 for more details about the sample, specifications, control variables, and outcomes. The outcome variable in sub-figure (a) is whether the older cousin has an ADHD diagnosis in the outpatient data, and the outcome variable in sub-figure (b) is whether the older cousin has at least one ADHD drug claim in the prescription drug data. The outcome variable in sub-figure (c) is whether the younger cousin has an ADHD diagnosis in the outpatient data, and the outcome variable in sub-figure (d) is whether the younger cousin has at least one ADHD drug claim in the prescription drug data.

Figure B8: Stylized Framework for Interpreting ADHD Gap; Three Possible Interpretations

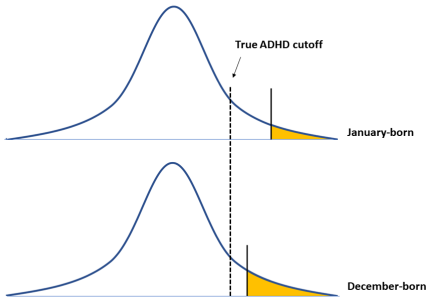
(a) Over-Diagnosis in December, Under-Diagnosis in January



(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 day of birth.

Table B1: Sample Means of Key Variables

Panel A: Older Cousins	Full Sample	Before Cutoff	After Cutoff
Share w/ ADHD diagnosis	0.036	0.038	0.034
Share w/ ADHD drug use	0.040	0.042	0.039
Father is foreign-born	0.061	0.064	0.059
Mother is foreign-born	0.053	0.056	0.051
Log household income	7.822	7.819	7.825
Father has college degree+	0.124	0.123	0.125
Mother has college degree+	0.105	0.104	0.106
Number of cousins	1.984	1.991	1.978
Observations	575,218	267,794	307,424
Panel B: Younger Cousins	Full Sample	Before Cutoff	After Cutoff
Share w/ ADHD diagnosis	0.043	0.043	0.042
Share w/ ADHD drug use	0.048	0.049	0.048
Father is foreign-born	0.059	0.061	0.058
Mother is foreign-born	0.053	0.055	0.052
Log household income	7.797	7.788	7.805
Father has college degree+	0.121	0.119	0.122
Mother has college degree+	0.110	0.107	0.112
Birth spacing (in months)	28.965	29.052	28.888
Observations	1,122,756	524,409	598,347

Notes: This table reports sample means of some of the variables in our analysis. The first column uses our full analysis sample of of cousin pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. The second and third columns split the sample into families with older cousins born before and after the cutoff (January 1st), respectively.

Table B2: Results for Placebo Outcomes

	(1) Gestation length	(2) Share boys	(3) Fathers' educ	(4) Mothers' educ	(5) Birth spacing
Panel A: Older Cousins					
OC born before the cutoff	-0.0777	0.0082*	-0.0008	-0.0019	
	(0.1310)	(0.0042)	(0.0025)	(0.0022)	
Mean(Y)	278.043	0.514	0.123	0.104	
N	221,542	221,542	221,542	221,542	
Panel B: Younger Cousins					
OC born before the cutoff	-0.1343	-0.0009	0.0008	-0.0002	0.1351
	(0.0992)	(0.0030)	(0.0024)	(0.0023)	(0.0848)
Mean(Y)	278.439	0.513	0.120	0.108	28.788
N	432,693	432,693	432,693	432,693	432,693

Notes: Each column reports results from a separate regression. The sample of analysis is the universe of cousins pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. In Panels (A) and (B), we report the placebo outcomes of the older and younger cousins, respectively. We use a bandwidth of 75 days and include the same full set of control variables as in our main specifications as in Tables 1 and 2, except we omit indicators for the older and younger cousin's gender in column (2) of Panels A and B. We omit the father's education categories in column (3) and the mother's education categories in column (4). We also omit the birth spacing between cousin pairs in column (5). Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B3: Effect of Older Cousin Being Born Before Cutoff on Own and Younger Cousin's ADHD Outcomes, Different Polynomials of the Running Variable

	ADHD Diag			ADHD Drug		
	(1) Linear	(2) Quadratic	(3) Cubic	(4) Linear	(5) Quadratic	(6) Cubic
Panel A: Older Cousins						
OC born before the cutoff	0.0087***	0.0069***	0.0060**	0.0085***	0.0066***	0.0060**
[Own Relative Age Effect]	(0.0012)	(0.0018)	(0.0025)	(0.0011)	(0.0017)	(0.0024)
Mean(Y)	0.036	0.036	0.036	0.040	0.040	0.040
N	575,218	575,218	575,218	575,218	575,218	575,218
Panel B: Younger Cousins						
OC born before the cutoff	0.0041***	0.0039***	0.0033**	0.0038***	0.0028**	0.0016
[Spillover Effect]	(0.0008)	(0.0011)	(0.0015)	(0.0009)	(0.0013)	(0.0016)
YC born before the cutoff	0.0125***	0.0119***	0.0125***	0.0133***	0.0117***	0.0122***
[Own Relative Age Effect]	(0.0009)	(0.0013)	(0.0018)	(0.0009)	(0.0014)	(0.0018)
Mean(Y)	0.043	0.043	0.043	0.048	0.048	0.048
N	1,122,756	1,122,756	1,122,756	1,122,756	1,122,756	1,122,756

Notes: Each column reports results from a separate regression. We use a global bandwidth and include the same set of control variables as in Panel B of Tables 1 and 2. See notes under Tables 1 and 2 for more details about the sample, specifications, control variables, and outcomes. In columns (2) and (4), we include quadratic polynomials of the running variables, and in columns (3) and (6), we include cubic polynomials of the running variables. Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B4: Effect of Older Cousin Being Born Before Cutoff on Own and Younger Cousin's ADHD Outcomes, "Doughnut-RD"

	(1) ADHD Diag	(2) ADHD Drug
Panel A: Older Cousins		
OC born before the cutoff	0.0061**	0.0073***
[Own Relative Age Effect]	(0.0029)	(0.0026)
Mean(Y)	0.037	0.041
N	183,016	183,016
Panel B: Younger Cousins		
OC born before the cutoff	0.0061***	0.0055**
[Spillover Effect]	(0.0020)	(0.0023)
YC born before the cutoff	0.0121***	0.0131***
[Own Relative Age Effect]	(0.0015)	(0.0016)
Mean(Y)	0.043	0.048
N	356,729	356,729

Notes: Each column reports results from a separate regression. We use a bandwidth of 75 days and include the same set of control variables as in Panel B of Tables 1 and 2. See notes under Tables 1 and 2 for more details about the sample, specifications, control variables, and outcomes. We additionally exclude all cousins pairs with older cousins born in the two-week bandwidth around the cutoff (January 1st). Robust standard errors are clustered on the older cousin's day of birth.
Significance levels: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table B5: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Outcomes, Non-Parametric RD Models

	(1) MSE	(2) MSE-2	(3) MSE-Sum	(4) Min-MSE	(5) Med-MSE	(6) CER	(7) CER-2	(8) CER-Sum	(9) Min-CER	(10) Med-CER
Panel A: ADHD Diagnoses										
OC born before the cutoff	0.0029	0.0029	0.0038*	0.0038*	0.0029	0.0082***	0.0075***	0.0091***	0.0091***	0.0089***
[Spillover Effect]	(0.0018)	(0.0018)	(0.0020)	(0.0020)	(0.0019)	(0.0026)	(0.0026)	(0.0028)	(0.0028)	(0.0027)
Mean(Y)	0.042	0.043	0.042	0.042	0.043	0.042	0.043	0.043	0.043	0.042
N	261,871	276,680	215,045	215,045	235,309	128,986	135,267	106,656	106,656	117,048
Left BW	46.24	33.43	39.00	39.00	39.00	23.04	16.66	19.43	19.43	19.43
Right BW	46.24	68.69	39.00	39.00	46.24	23.04	34.23	19.43	19.43	23.04
Panel B: ADHD Drug Treatment										
OC born before the cutoff	0.0010	0.0018	0.0018	0.0018	0.0012	0.0063**	0.0058**	0.0070**	0.0070**	0.0068**
[Spillover Effect]	(0.0019)	(0.0019)	(0.0021)	(0.0021)	(0.0020)	(0.0028)	(0.0027)	(0.0029)	(0.0029)	(0.0029)
Mean(Y)	0.048	0.049	0.048	0.048	0.048	0.048	0.049	0.048	0.048	0.048
N	244,413	279,432	215,045	215,045	227,790	118,043	135,267	106,656	106,656	111,938
Left BW	43.98	33.73	38.92	38.92	38.92	21.92	16.81	19.40	19.40	19.40
Right BW	43.98	69.19	38.92	38.92	43.98	21.92	34.48	19.40	19.40	21.92

Notes: Each column reports results from a separate regression. The sample and outcomes are the same as in Table 2. Each column shows results from an RD model with local linear polynomials, triangular kernels, and robust bias-corrected inference procedures, using different optimal bandwidth algorithms to select the bandwidths of the number of days used on each side of the cutoff in the older cousin's date of birth relative to the school entry cutoff. Panel A shows results using the younger cousin's ADHD diagnosis as the outcome, while Panel B shows results using the younger cousin's ADHD drug treatment as the outcome. The optimal bandwidth algorithms are: (1) one common mean squared error (MSE)-optimal bandwidth selector for both sides of the cutoff; (2) two different MSE-optimal bandwidth selectors (below and above the cutoff); (3) one common MSE-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof); (4) minimum of (1) and (3); (5) median of (1), (2), and (3) for each side of the cutoff separately; (6) one common coverage error rate (CER)-optimal bandwidth selector; (7) two different CER-optimal bandwidth selectors (below and above the cutoff); (8) one common CER-optimal bandwidth selector for the sum of regression estimates (as opposed to difference thereof); (9) minimum of (6) and (8); (10) median of (6), (7), and (8) for each side of the cutoff separately. We use the Stata "rdrobust" command for these analyses (Calonico et al., 2017). We report the number of days used in the left and right-hand bandwidths in each model at the bottom of the table. All regressions control for the same set of controls as in Table 2, as well as a linear spline function of the younger cousin's own date of birth relative to the cutoff. Robust standard errors are reported in parentheses.

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