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

Petra Persson† Xinyao Qiu‡ Maya Rossin-Slater§

May 27, 2024

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 Hofmann and Joshua Bricker, as well as Sarah Bögl and Iliriana Shala 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 relatives. 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 relatives are 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 (Bos, Hertzberg and Liberman, 2020; Cuddy and Currie, 2020; Alalouf, Miller and Wherry, 2024), while a 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 percent of children in the United States (Danielson et al., 2018) and seven percent of children

[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, Salz and Lillie, 2007; Bond et al., 2013; Ong and Mandl, 2015; Einav et al., 2020) and pneumonia (Chan, Gentzkow and Yu, 2022). 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."

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 those who are the oldest. 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 17.6 percent more likely to be diagnosed with ADHD and 16.1 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 multiple channels (Landersø, Nielsen and Simonsen, 2019; Karbownik and Özek, 2023), we use *non-nuclear* family members—first cousins—to study spillovers of relative-age-induced marginal ADHD diagnoses.[7] Specifically, we show that younger cousins of children born just before the school entry cutoff are 9.3 and 6.0 percent more likely to be diagnosed with ADHD and treated with ADHD drugs, respectively, than the

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

[5]See, for example: Elder, 2010; Evans, Morrill and Parente, 2010; Dalsgaard et al., 2012; Morrow et al., 2012; Zoëga, Valdimarsdóttir and Hernández-Díaz, 2012; Halldner et al., 2014; Krabbe et al., 2014; Pottegård, Hallas and Zoëga, 2014; Chen et al., 2016; Schwandt and Wuppermann, 2016; Layton et al., 2018; Whitely et al., 2018; Root et al., 2019; Furzer, 2020; Furzer, Dhuey and Laporte, 2022.

[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 et al., 2014; Furzer, Dhuey and Laporte, 2022), and demonstrate heterogeneity in the types of diagnostic errors with respect to child gender, race, and socioeconomic status (Furzer, 2020; Marquardt, 2020).

[7]In our data, we show that a child's relative age for grade influences their own mother's labor market and marital outcomes, but does not influence the outcomes of their younger cousin's mother (i.e., their aunt). See Section 4.1 for more details.

younger cousins of children born shortly after the cutoff. Importantly, these discontinuities exist *conditional on the younger children's own relative age for grade.* In fact, the magnitudes of the estimated spillover effects on cousins amount to around one-third and one-fifth of the younger cousins' own relative age effects on ADHD diagnoses and drug treatment, respectively. Scaling by the older cousins' own relative age effects suggests that each marginal diagnosis among older cousins generates 0.6 additional diagnoses among younger cousins. Analysis using information on diagnosis timing provides further support for the spillover: 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.

To shed light on whether these ADHD diagnosis spillovers confer any long-term benefits, we analyze several educational and labor market outcomes: high school grade point average (GPA), an indicator for on-time high school graduation, an indicator for college enrollment by age 21, and average annual earnings between ages 27 and 30. We do not find that the 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 significantly better outcomes than the younger cousins of those who are born after the cutoff. Indeed, the (imprecisely estimated) coefficients in our preferred specifications are all negatively signed.[8] Overall, these results suggest that the younger cousins of marginally diagnosed children are no better off in the long-term; if anything, they may even be worse off.

Why would the younger cousin of a marginally diagnosed child be more likely to be diagnosed with and treated for ADHD? To shed light on the mechanisms behind these spillovers, we consider the key steps in the ADHD diagnosis process: first, one must request an ADHD evaluation that can lead to a referral for a screening, and second, a physician performs an ADHD screening. As we describe in Section 2, both families and schools can request an ADHD evaluation in Sweden. Since most cousins do not attend the same schools,[9] intra-

[8]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, Devereux and Salvanes, 2011; Kawaguchi, 2011; Fredriksson and Öckert, 2014; Hurwitz, Smith and Howell, 2015; Depew and Eren, 2016; Cook and Kang, 2016; Landersø, Nielsen and Simonsen, 2017; Dhuey et al., 2019).

[9]We do not have information about which schools children in our data attend, and thus cannot calculate

family communication between parents—who are siblings themselves—is likely a key channel. We find that the ADHD spillovers are stronger among cousin pairs in which at least one of the sibling parents is a woman, which is consistent with a body of evidence showing that women are more likely to communicate family health information than men (e.g., Wilson et al., 2004; Koehly et al., 2009; Montgomery et al., 2013; Mendes et al., 2016).

The intra-family communication channel implies that the younger cousins of marginally diagnosed older children are more likely to end up at a physician's office for an ADHD screening than the younger cousins of those who are not. At the screening stage, the physician's diagnostic technology plays a critical role. For health conditions with a precise diagnostic technology (e.g., a genetic test), this "evaluation-seeking gap" would not necessarily translate into a difference in diagnosis rates. However, ADHD falls into a large class of health conditions for which there is no precise technology to rule out erroneous or low-value diagnoses; instead, physicians have a noisy screening protocol. Moreover, the ADHD screening protocol indicates that family history of ADHD—including among cousins—is considered, but does *not* account for the likely severity of a case based on the previously diagnosed family member's relative age for grade. Consistent with this, we show the existence of ADHD spillovers on different levels of treatment intensity among younger cousins—we find effects on cases in which younger cousins are diagnosed and treated with drugs, as well as on cases in which a diagnosis is not followed by any drug treatment. We additionally find an increase in the number of outpatient mental health-related visits in the three years following the initial diagnosis. All in all, it appears that physicians do not undo the "evaluation-seeking gap," such that it translates into a "diagnosis gap" that in turn leads to differences in the utilization of ADHD-related healthcare services.

In sum, our findings suggest that marginal ADHD diagnoses propagate across cousins, and there is no clear evidence that these spillovers result in human capital gains in the longer run. Using our estimates of spillovers on the number of ADHD drugs and mental health outpatient visits as well as the average costs of these treatments, we estimate that each diagnosis spillover raises healthcare costs in Sweden by about \$1,565 (in 2019 USD). This cost is likely to be higher in countries with higher prices for ADHD drugs and outpatient therapy, like the United

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 exists in the sample of cousin pairs who live in different municipalities.

States.

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, Salz and Lillie, 2007; Bond et al., 2013; Ong and Mandl, 2015; Mullainathan and Obermeyer, 2017; Obermeyer et al., 2019; Einav et al., 2020; Chan, Gentzkow and Yu, 2022; Alalouf, Miller and Wherry, 2024). Particularly relevant to our study of a mental health condition, Bos, Hertzberg and Liberman (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 *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); Cuddy and Currie (2020); Currie and MacLeod (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, Currie and Sonchak, 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. Therefore, spillover effects may be an important mechanism underlying the "exploding" ADHD caseloads phenomenon (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' human capital and health outcomes (see, e.g., Fletcher, Hair and Wolfe, 2012; Breining, 2014; Parman, 2015; Yi

[10]For evidence on sibling spillovers in test scores, educational attainment, college choice, and major choice, see, e.g., Dustan (2018); Joensen and Nielsen (2018); Qureshi (2018*a*,*b*); Nicoletti and Rabe (2019); Dahl, Rooth and Stenberg (2020); Aguirre and Matta (2021); Altmejd et al. (2021); Karbownik and Özek (2023). For sibling spillovers in military service, see Bingley, Lundborg and Lyk-Jensen (2019); for spillovers in program take-up (e.g., paternity leave), see Dahl, Løken and Mogstad (2014).

et al., 2015; Alsan, 2017; Black et al., 2021; Daysal et al., 2022, 2021), arguing that shifts in parental resource allocation across siblings and within-family infectious disease spread may be important mechanisms.[11] Our paper provides novel evidence of health-related spillovers across non-nuclear family members—cousins—and also relates to Chen, Persson and Polyakova (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: initial evaluation, referral for a formal screening, and diagnosis.

Initial evaluation and referral for ADHD screening. In order to be diagnosed with ADHD, an individual needs to go through a neuropsychiatric screening (in Swedish: *Neuropsykiatrisk utredning*).[14] For simplicity, we refer to it as an "ADHD screening" throughout.

A referral for an ADHD screening is made after an initial evaluation, which can be requested by a parent or by a school.[15] The initial evaluation assesses the needs of the child, whether

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

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

[13] Under publicly provided universal health insurance, 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] This evaluation may lead to an ADHD diagnosis or to other neuropsychiatric diagnoses such as autism or Tourette's syndrome (or to no diagnosis). See 1177 Vårdguiden (2022).

[15] All children 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).

referral should be made for a full ADHD screening, and whether the child needs any support in the interim period, while waiting for such a screening (1177 Vårdguiden, 2022).[16] This initial evaluation can be conducted by several entities in the Swedish healthcare system: a local primary care center (in Swedish: *vårdcentral*), a pediatrician, the child's school health care system, or a pediatric psychiatrist (in Swedish: *Barn- och ungdomspsykiatrin* (1177 Vårdguiden, 2022).

ADHD screening and diagnosis. An ADHD screening involves several components, and is similar across many countries, including Sweden and the United States.[17] First, using information from interviews with parents and teachers or other caregivers, the child is assessed with 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.[18]

Second, the ADHD screening includes a physical exam and an evaluation of the child's family history of ADHD. In the US, the CDC specifies that the healthcare provider should be given information about the health history of a broad group of relatives, including the child's siblings, parents, grandparents, aunts and uncles, and *cousins*.[19] The protocols do not specify more precisely *how* the child's family history of ADHD should be incorporated into

[16] While treatment with ADHD medication requires a formal ADHD diagnosis (and thus a full ADHD sceening), other accommodations, such as extra support at school, can be made after the initial assessment (Socialstyrelsen, 2022).

[17] The process in Sweden is described in detail in Socialstyrelsen (2014) and the process in the United States is described by the Centers for Disease Control and Prevention (CDC) and on *UpToDate*, a service that aggregates medical research for clinical practice. See: `https://www.cdc.gov/ncbddd/adhd/diagnosis.html`, accessed on February 15, 2024, and `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.

[18] 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.

[19] The exact formulation is "Collect your child's family health history information before seeing your child's health care provider: Include your and your partner's children, parents, sisters, brothers, grandparents, aunts, uncles, nieces, and nephews." See: `https://www.cdc.gov/genomics/disease/attention.htm`, accessed on February 15, 2024. In Sweden, while we have not found precise information about cousins in any formal written guidelines that mention "family history," we have confirmed with a Swedish psychologist who conducts ADHD screenings that information about cousins is used if families bring it up.

the diagnostic process.[20]

A child interacts with several types of healthcare providers during the screening process, and Sweden's national guidelines stipulate that each screening should be individualized, but generally overseen by at least one psychiatrist and one psychologist (Socialstyrelsen, 2022).

ADHD treatment. Children diagnosed with ADHD can be treated with pharmaceutical drugs, psychotherapy, or with a combination of both. 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 under age 18.[21] 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 or pediatric neurologists in Sweden.

Figure 1 plots the trend in ADHD diagnoses and drug treatment rates among Swedish-born 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 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.[22] 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

[20]For example, *UpToDate* 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.

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

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

is around 21 years.[23]

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.[24]

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 set with first cousin identifiers. We then construct pairs of cousins, where each pair has an older and a younger child.[25] We keep pairs with an age difference of no more than five 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.[26] Our final analytic sample consists of 1,122,772 cousin pairs,

[23] 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).

[24] 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.

[25] 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.

[26] Our analysis essentially aggregates RDs across 11 fiscal years of birth of the older child, from July 1985–

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).[27] 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.[28]

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 the younger cousin is born in 1992 or earlier, and calculate average annual earnings over ages 27–30.[29]

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.[30]

June 1986 through July 1995–June 1996. We include fixed effects for the older child's fiscal year of birth in our regression models.

[27]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).

[28]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.

[29]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.

[30]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

Sample means. Appendix Table C1 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 and January-June, 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 the small differences are statistically significant due to our large sample size and the fact that this table uses data on births throughout the entire year. These differences further motivate our RD approach, which only uses data from a narrow bandwidth around the school entry cutoff.

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 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.[31] Our primary specifications use a bandwidth of 75 days around the cutoff and a linear spline parametrization; in Appendix C, we explore the sensitivity of our estimates to other bandwidths and higher-

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.

[31]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).

order polynomials, as well as to non-parametric RD models with optimal bandwidth selection algorithms (Calonico, Cattaneo and Titiunik, 2014*a*,*b*; Calonico et al., 2017, 2019).

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

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

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 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.

4.1 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.

Non-random sorting. To assess issue (1), we begin by plotting a histogram of births at a daily level using our sample of older cousins in Figure 2 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, Jansson and Ma, 2018) yields a statistically significant $t-$statistic of 8.11.

The evidence of non-random sorting of births is consistent with a number of studies from other contexts that document that parents are able to manipulate the timing of childbirth in order to, for example, qualify for financial benefits (Dickert-Conlin and Chandra, 1999; Gans and Leigh, 2009; Buckles and Hungerman, 2013; Schulkind and Shapiro, 2014; LaLumia, Sallee and Turner, 2015). Notably, Dickert-Conlin and Elder (2010) emphasize that birth timing manipulation around school-entry cutoffs only tends to occur when these cutoffs coincide with holidays, as is the case in our setting of a cutoff on New Year's Day. To investigate the nature of this manipulation, Appendix Figure C1 plots the number of births by week of the year in our sample of older cousins separately for those that are delivered via a planned cesarean section (sub-figure a), those that are induced (sub-figure b), and those that are non-induced vaginal deliveries (sub-figure c). There is a clear dip in the distribution of planned c-sections and inductions in the last week of December, consistent with physicians avoiding scheduling such procedures during the holiday season. We do not see such a dip for non-induced vaginal deliveries, which are much less susceptible to timing manipulation.

Importantly, the significant difference in the number of births between the end of December and the beginning of January only poses a concern for our analysis if the sorting is systematically related to our outcomes of interest. To assess this possibility, we examine whether there are any differences in birth-related and 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 the average gestation length (in days) of the older or younger cousins (Appendix Figure C2), average birth spacing between cousins (Appendix Figure C3), the gender composition of older and younger cousins (Appendix Figure C4), or parental education levels (Appendix Figure C5). Appendix Figure C6 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.[32] We do not see any discontinuities in these predicted outcomes at the school entry

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

cutoff. Lastly, Appendix Table C2 reports results from RD regressions that use the birth and family background characteristics as outcomes (Pei, Pischke and Schwandt, 2019); we find no evidence of statistically significant (at the 5% level) discontinuities based on the older cousin being born before versus after the January 1 cutoff.

In sum, this analysis does not reveal any systematic discontinuities in observable characteristics at the school-entry cutoff, suggesting that the sorting observed in Figure 2 is unlikely to bias our RD design. To further address the concern about sorting, we conduct two additional robustness checks that are reported in Appendix C. First, we estimate a "doughnut-RD" model that omits all cousin pairs with older cousins born in a two-week bandwidth surrounding the cutoff. Second, we use a sub-sample of cousin pairs in which the older cousins were born via non-induced vaginal deliveries, as the timing of these deliveries is difficult to manipulate. Our results are robust to both of these sample restrictions.

Potential alternative channels. 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, Nielsen and Simonsen (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 (2023) 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.[33]

Importantly, these family-level dynamics have only been documented within nuclear families—i.e., for siblings who share the same parents. We investigate the extent to which such responses are mitigated when we consider cousins, who do not share the same parents (and almost never share the same household in the Swedish context). Appendix Tables C3 and C4 present results from estimating models (1) and (2) using as dependent variables the labor market and marital outcomes of the mothers of older and younger cousins, respectively. Consistent with Landersø, Nielsen and Simonsen (2019)'s evidence from Denmark, we find that an older

[33] See also Altmejd et al. (2021) for evidence of sibling spillovers in college enrollment and college choice.

cousin's relative age for grade affects his/her own mother's employment, work income, and marital status measured when the child is 7 years old. However, we do not see any evidence of spillover effects on the younger cousin's mother's labor market or marital outcomes. These analyses provide some reassurance that studying ADHD diagnosis spillovers across cousins is a way of assuaging concerns about alternative channels, although of course it is impossible to definitively rule out all other possible explanations.

5 Results

We begin by using our analysis sample of older cousins to confirm prior evidence (e.g., Elder, 2010; Evans, Morrill and Parente, 2010; Dalsgaard et al., 2012; Morrow et al., 2012; Zoëga, Valdimarsdóttir and Hernández-Díaz, 2012; Halldner et al., 2014; Krabbe et al., 2014; Pottegård, Hallas and Zoëga, 2014; Chen et al., 2016; Schwandt and Wuppermann, 2016; Layton et al., 2018; Whitely et al., 2018; Root et al., 2019; Furzer, Dhuey and Laporte, 2022) 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 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 3 plots raw data means of the 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. Focusing on the estimates from models with controls in Panel B, 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 17.6 and 16.1 percent, respectively.

Panel A of Appendix Table C5 shows the sensitivity of our estimates to using different polynomials in the running variable, while sub-figures (a) and (b) of Appendix Figure C7 show results using different bandwidths. In addition, to address the issue of potential non-random sorting around the threshold, Appendix Table C6 shows results from a restricted sample of older cousins born via a non-induced vaginal delivery, while Panel A of Appendix Table C7 presents results from "doughnut-RD" models, which omit cousins born in a two-week bandwidth surrounding the cutoff. 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.

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 3, except that it plots raw data means of the younger cousins' ADHD diagnosis and drug treatment rates on the respective sub-figure $y-$axes. It appears 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 cousins. 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 B (with controls) show that younger cousins born before the school entry cutoff are 1.2 and 1.3 percentage points more likely to have an ADHD diagnosis and an ADHD drug claim, respectively, than those born after the cutoff. Notably, the younger cousins' own age-for-grade effects on ADHD outcomes are about twice as large in magnitude than the older cousins' own age-for-grade effects reported in Table 1. The larger own age-for-grade effect magnitudes among younger cousins are consistent with an overall increasing trend in ADHD diagnosis and drug treatment rates over time.

Further, conditional on the younger cousin's own relative age for grade, we find a significant effect of the older cousin's relative age for grade on the younger cousin's likelihood of ADHD diagnosis and drug treatment. The estimates presented in Panel B show that 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 9.1 and 5.8 percent effect sizes when evaluated at the respective dependent variable means. The magnitudes of the spillovers on ADHD diagnoses and drugs are 32.8 and 21.4 percent, respectively, of the sizes of the younger cousin's own relative age-for-grade effects on these outcomes. Scaling by the own relative age-for-grade effects among the older cousins reported in Table 1 implies that each marginal diagnosis among older cousins leads to 0.6 additional diagnoses among the younger cousins. The magnitudes of the spillover effects in ADHD outcomes relative to own age-for-grade effects are slightly smaller than those found for educational outcomes among siblings—for example, Karbownik and Özek (2023) show that the sibling spillover effect on test scores is about two-thirds of the own age-for-grade effect on this outcome.

Panel B of Appendix Table C5 shows the sensitivity of our spillover estimates to using different polynomials in the running variables, while sub-figures (c) and (d) of Appendix Figure C7 present the spillover results based on RD models with different bandwidths. Additionally, Panel B of Appendix Table C7 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, while Appendix Table C8 presents results from a sub-sample in which the older cousins are born via a non-induced vaginal delivery. Finally, Appendix Table C9 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.[34]

These analyses indicate that our ADHD spillover results are not sensitive to using bandwidths in the range of 55 to 100 days around the cutoff date (Appendix Figure C7) nor to the specifications that address concerns about non-random sorting around the threshold (Appendix Tables C7 and C8). Appendix Table C9 shows that our results are robust to using variations of the coverage error rate (CER) optimal bandwidth selector, which yield bandwidths of 16–23 days on the left and 19–34 days on the right. The estimates are weaker when we use various mean squared error (MSE) optimal bandwidth selectors, which yield wider bandwidths that might necessitate higher order polynomials in the running variable. At the same time, when we use a global bandwidth, we find that our results are not sensitive to using linear, quadratic, or cubic polynomials (Appendix Table C5).

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 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

[34]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.

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), Panel B, 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 0.4 percentage points higher relative to a case where the older cousin is born after the cutoff. Conversely, in column (3), Panel B, 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 0.5 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

Do the spillovers of marginal ADHD diagnoses translate into any long-term benefits for the younger cousins? To shed light on this question, we study our three educational outcomes and annual earnings averaged over ages 27–30. Figure 5 presents raw data means of these outcomes observed among the younger cousins by the older cousin's week of birth, while Table 4 presents results from estimating equation (2), with and without family background controls.

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, Devereux and Salvanes, 2011; Kawaguchi, 2011; Fredriksson and Öckert, 2014; Hurwitz, Smith and Howell, 2015; Depew and Eren, 2016; Cook and Kang, 2016; Landersø, Nielsen and Simonsen, 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 do not appear to have any better outcomes than the younger cousins of children born after the cutoff. The results from specifications with control variables in Panel B are not statistically significant at the 5% level for three out of the four outcomes analyzed. While there does appear to be a significant negative effect on the younger cousin's high school GPA, it is not robust across our sensitivity analyses (see Appendix Figure C8 and Appendix Table C10), and we therefore suggest caution for interpreting its exact magnitude. Overall, these findings imply that the long-run educational and economic outcomes among younger cousins are not any better among those who are marginally diagnosed with ADHD; if anything, they may potentially be worse.

These results additionally have implications for understanding the consequences of ADHD drug treatment. As we discuss below in Section 5.4, we find spillover effects on several measures of drug consumption, including persistence of use over time and the number of unique drugs prescribed. This is consistent with other evidence that a positive ADHD diagnosis can lead to long-term use of ADHD prescription drugs—50 percent of patients who initiate ADHD drugs remain on them five years later (Socialstyrelsen, 2012).[35] Existing research on the impacts of using ADHD drugs is mixed and limited to studies of selected short- and medium-term behavioral and educational outcomes.[36] To date, we know very little about the longer-term impacts of ADHD drug treatment on measures of individuals' well-being. Our estimates suggest that there are no clear long-term human capital gains associated with the use of ADHD drugs on the margin (and that there could even be some costs).

5.4 Potential Mechanisms Behind Marginal ADHD Spillovers

We have established that there exist spillovers of marginal ADHD diagnoses across cousins, and that these spillovers do not result in any long-term human capital benefits (and may even lead to some costs). In this section, we discuss the potential mechanisms that underlie these spillover effects.

To interpret this spillover, it is first helpful to consider the etiology of ADHD. Unlike con-

[35] Appendix Table C11 shows that this is also true in our sample. Specifically, we use all individuals in our baseline sample who receive an ADHD diagnosis after July 2005 (when the prescription drug data begin), and who obtain an ADHD drug prescription within one year of the diagnosis. We then report the shares of these individuals who also have an ADHD drug claim two, three, four, and five years later. We find that 48 percent of individuals still have an ADHD drug claim by the end of the 4th year following a diagnosis.

[36] See, e.g., Jensen, 1999; Wilens et al., 2003; Charach, Ickowicz and Schachar, 2004; Dalsgaard et al., 2012; Humphreys, Eng and Lee, 2013; Molina et al., 2013; Currie, Stabile and Jones, 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, Pai and Suziedelyte, 2017).

ditions 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. 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.

Appendix Figure C9 visualizes these ideas graphically to aid the interpretation of the observed difference in ADHD diagnoses between the December- and January-born older cousins (Figure 3 and Table 1). The bell curves represent the distributions of underlying ADHD symptoms in the populations of children born in December and January, respectively. The yellow areas under each of the curves signify the children who receive positive ADHD diagnoses. There are more diagnoses among December-born than January-born children. Depending on the location of the "true" ADHD cutoff, one can interpret this pattern as a combination of over-diagnoses among December-born and under-diagnoses among January-born children (Appendix Figure C9(b)), more over-diagnoses among December-born than January-born children (Appendix Figure C9(c)), or fewer under-diagnoses among December-born than January-born children (Appendix Figure C9(d)). Since the "true" cutoff is unobserved in our data (or any other data set available to researchers), we do not take a stance on the direction of error represented by the diagnosis gap.[37] However, a key takeaway is that, regardless of the interpretation, *the average diagnosed December-born child has lighter symptoms than the average diagnosed January-born child.*

With this interpretation of the diagnosis gap among older cousins in mind, what can we say

[37] 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.

about how it might propagate to their younger cousins? As described in Section 2, the ADHD diagnosis process involves several key steps: First, one must seek out an ADHD evaluation that can lead to a referral for a screening, and second, a physician performs a screening (which may or may not result in a diagnosis). The existence of spillovers means that somewhere in this process, a marginal diagnosis of an older cousin affects the likelihood that a younger cousin is diagnosed.

Evaluation-seeking stage. A spillover at the evaluation-seeking stage would indicate that the families of younger cousins of marginally diagnosed older children are more likely to request ADHD evaluations than the families of younger cousins of older children who are not. This seems quite plausible since the parents are siblings, who likely communicate about their children's health issues. Put differently, the parent of a marginally diagnosed older cousin is likely to share information about their child's ADHD diagnosis with their sibling, who may then, in turn, be more likely to seek out an evaluation for their own child (i.e., the younger cousin). This may occur because the parents of the younger cousin learn about advantages of an ADHD diagnosis (e.g., receiving extra time on tests at school) or receive information about the consequences of drug treatment (e.g., better focus).

Since women are more likely to communicate family health information than men (see, e.g., Wilson et al., 2004; Koehly et al., 2009; Montgomery et al., 2013; Mendes et al., 2016), we investigate the importance of intra-family communication as a channel by studying heterogeneity in ADHD spillovers across cousin pairs in which the sibling parents are both mothers (i.e., sisters), both fathers (i.e., brothers), or a mother and a father (i.e., a sister and a brother). Results from this analysis are presented in Appendix Table C12. We find that spillovers only exist among cousin pairs in which at least one of the sibling parents is a mother, whereas there are no spillovers when the related parents are both fathers. Moreover, compared to the younger cousin's own relative age effect, the magnitude of the spillover is largest in cousin pairs in which both sibling parents are mothers (i.e., sisters). These patterns suggest that intra-family communication about children's diagnoses—perpetuated by adult sisters—is a likely mechanism driving the ADHD spillovers.

We have also examined heterogeneity in the ADHD spillover effects across cousins by the income, foreign-born status, and educational attainment of the parents of the older cousins in Appendix Tables C13, C14, and C15, respectively. The coefficient magnitudes of the spillover effects are larger in cousin pairs in which the older cousin's household has below-median income, has a Swedish-born mother, and has a mother with no college education. Notably, the own relative age effects among younger cousins on ADHD diagnoses and drug take-up are fairly similar across these sub-groups. Thus, it appears that the spillover effects are somewhat stronger in less advantaged (but native Swedish) families, further supporting intra-family communication as an important channel, as shown in prior work. For example, Chen, Persson and Polyakova (2022) find that the effect of having a healthcare provider in the family on health behaviors is stronger in lower socioeconomic status families in Sweden, which is consistent with these families having less other sources of medical information compared to their more advantaged counterparts.

As noted in Section 2, schools may also play a role in requesting ADHD evaluations. However, given that the vast majority of cousins do not attend the same schools in our setting, this seems like a less plausible mechanism for driving the spillover at the evaluation-seeking stage.

Diagnosis stage. The "evaluation-seeking gap" just discussed implies that a higher share of children with an older cousin born just before the school entry cutoff end up in the doctor's office for an ADHD screening than of 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, 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.

However, ADHD falls into a large class of health conditions for which there is no discrete diagnostic test 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, the "evaluation-seeking gap" could translate into spillovers of marginal low-value—or potentially inaccurate—diagnoses.

Recall that Appendix Figure C9 implies that December-born older cousins who are diagnosed with ADHD should on average have lighter symptoms than their January-born counterparts. If physicians took this fact into account, then they might treat the younger cousins of December-born children as more marginal cases compared to the younger cousins of January-borns. We empirically investigate this conjecture using several "intensive margin" measures of ADHD treatment as outcomes among the younger cousins: (i) the persistence of ADHD drug treatment one to three years following diagnosis, (ii) the number of unique drugs used during treatment in the year after diagnosis, and (iii) the number of outpatient mental health-related visits one to three years following diagnosis.

Appendix Tables C16 and C17 present the results from these analyses. It appears that the spillover effect on drug treatment among younger cousins is stronger in the year immediately following diagnosis, than in years two and three (columns (1)-(3) of Appendix Table C16), although the coefficients across the three outcomes are not statistically different from each other. At the same time, column (4) of Appendix Table C16, indicates that there is also a spillover on the likelihood that the younger cousin is diagnosed with ADHD and is *not* taking any ADHD medication in the three years that follow, which could signal a less severe case.[38] We also find a spillover effect on the total number of unique ADHD drugs used by younger cousins (Appendix Table C17, column (1)), which could either reflect a more intensive treatment protocol or the need to experiment with multiple drugs to find one that works. And we see persistent increases in the number of mental health-related outpatient visits in the three years following the diagnosis (columns (2)–(4) of Appendix Table C17).

In sum, it appears that while some of the spillovers generate potentially less severe diagnoses that do not lead to medication use, we also observe spillovers of ADHD diagnoses that are accompanied by drug use and by extended outpatient mental health care. This is consistent with the fact that the physicians' ADHD screening protocol indicates that the presence of

[38] Appendix Table C16 reports that only 0.6 percent of individuals in our sample fall into the category of having an ADHD diagnosis but no medication in the three subsequent years. This share is low because we also include those who are not diagnosed with ADHD when calculating it. Conditional on having an ADHD diagnosis, about 14 percent of individuals in our analysis sample are never observed having any ADHD medication.

ADHD in the family is taken into account in the diagnostic process, but does not prescribe any specificity about how this information is used. Our analysis suggests that some physicians may recognize that the extra diagnoses among younger cousins are less severe cases, but others may not. Overall, however, it is clear that physicians do not fully "undo" the "evaluation-seeking gap" in ADHD, and that it instead translates into a diagnosis gap that generates differences in the utilization of ADHD-related healthcare services.

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 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 diagnosed 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 17.6 and 16.1 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 9.1 and 5.8 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 no better 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 some costs of ADHD diagnoses induced by the marginal diagnoses of older cousins.

We argue that intra-family communication likely plays an important role in the "tagging process" at the evaluation-seeking stage, in which younger cousins of previously diagnosed children are systematically more likely to seek out ADHD evaluations and be referred to ADHD screenings. Moreover, although physicians follow protocol by incorporating information about family history of ADHD in their diagnostic criteria, they do not undo the "evaluation-seeking gap" because they do not take into account the older child's relative age for grade and the implied severity of the case.

A back-of-the-envelope calculation suggests that each additional ADHD diagnosis spillover increases healthcare spending by $1,565 (in 2019 dollars) in terms of drug and outpatient treatment costs over the first three years following the spillover diagnosis.[39] A comparison to the social cost of treatment for an average ADHD patient suggests that a spillover diagnosis induces spending of approximately half of a typical ADHD patient in the first year following diagnosis, and at least one fifth of the cost over the first three years.[40] Thus, while spillover diagnosis cases likely have lighter symptoms on average, they nonetheless account for a substantial share of total spending on ADHD treatment.

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

[39] The reported costs are the sum of the costs borne by the public insurer as well as the (small) private out-of-pocket costs. We follow Deshpande and Mueller-Smith (2022) in using "intensive margin" outcomes when calculating costs. Specifically, we use the intensive margin estimates of the increased utilization of ADHD drugs and psychotherapy reported in Appendix Table C17.

[40] See Appendix B for details and the full cost calculation.

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

1177 Vårdguiden. 2022. "1177 Vårdguiden. Neuropsykiatrisk Utredning." Accessed February 2024. `www.1177.se/undersokning-behandling/undersokningar-och-provtagning/psykiatriska-utredningar/neuropsykiatrisk-utredning/`.

Aguirre, Josefa, and Juan Matta. 2021. "Walking in Your Footsteps: Sibling Spillovers in Higher Education Choices." *Economics of Education Review*, 80: 102062.

Alalouf, Mattan, Sarah Miller, and Laura R Wherry. 2024. "What Difference Does a Diagnosis Make? Evidence from Marginal Patients." *American Journal of Health Economics*, 10(1): 97–131.

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

Altmejd, Adam, Andrés Barrios-Fernández, Marin Drlje, Joshua Goodman, Michael Hurwitz, Dejan Kovac, Christine Mulhern, Christopher Neilson, and Jonathan Smith. 2021. "O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries." *The Quarterly Journal of Economics*, 136(3): 1831–1886.

Altonji, Joseph G., Sarah Cattan, and Iain Ware. 2017. "Identifying Sibling Influence on Teenage Substance Use." *Journal of Human Resources*, 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. 2018. "Analysis of Shared Heritability in Common Disorders of the Brain." *Science*, 360(6395): eaap8757.

Barkley, Russell A. 2006. "Attention-Deficit/Hyperactivity Disorder." In *Behavioral and Emotional Disorders in Adolescents: Nature, Assessment, and Treatment.* , ed. D.A. Wolfe and E.J. Mash. Guilford Publications.

Bedard, Kelly, and Elizabeth Dhuey. 2006. "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects." *The Quarterly Journal of Economics*, 121(4): 1437–1472.

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

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

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." *The Review of Economics and Statistics*, 93(2): 455–467.

Black, Sandra E, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen. 2021. "Sibling Spillovers." *The Economic Journal*, 131(633): 101–128.

Bond, Mary, Toby Pavey, Karen Welch, Chris Cooper, Ruth Garside, Sarah Dean, and Christopher J Hyde. 2013. "Psychological Consequences of False-Positive Screening Mammograms in the UK." *BMJ Evidence-Based Medicine*, 18(2): 54–61.

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

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

Brewer, Noel T, Talya Salz, and Sarah E Lillie. 2007. "Systematic Review: The Long-Term Effects of False-Positive Mammograms." *Annals of Internal Medicine*, 146(7): 502–510.

Buckles, Kasey S, and Daniel M Hungerman. 2013. "Season of Birth and Later Outcomes: Old Questions, New Answers." *Review of Economics and Statistics*, 95(3): 711–724.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014*a*. "Robust Data-Driven Inference in the Regression-Discontinuity Design." *The Stata Journal*, 14(4): 909–946.

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

Calonico, Sebastian, Matias D Cattaneo, Max H Farrell, and Rocio Titiunik. 2017. "rdrobust: Software for Regression-Discontinuity Designs." *The Stata Journal*, 17(2): 372–404.

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

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2018. "Manipulation Testing Based on Density Discontinuity." *The Stata Journal*, 18(1): 234–261.

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

Charach, Alice, Abel Ickowicz, and Russell Schachar. 2004. "Stimulant Treatment over Five Years: Adherence, Effectiveness, and Adverse Effects." *Journal of the American Academy of Child & Adolescent Psychiatry*, 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. 2016. "Influence of Relative Age on Diagnosis and Treatment of Attention-Deficit Hyperactivity Disorder in Taiwanese Children." *The Journal of Pediatrics*, 172: 162–167.

Chen, Yiqun, Petra Persson, and Maria Polyakova. 2022. "The Roots of Health Inequality and the Value of Intra-Family Expertise." *American Economic Journal: Applied Economics*.

Chorniy, Anna, and Leah Kitashima. 2016. "Sex, Drugs, and ADHD: The Effects of ADHD Pharmacological Treatment on Teens' Risky Behaviors." *Labour Economics*, 43: 87–105.

Chorniy, Anna, Janet Currie, and Lyudmyla Sonchak. 2018. "Exploding Asthma and ADHD Caseloads: The Role of Medicaid Managed Care." *Journal of Health Economics*, 60: 1–15.

Cook, Philip J, and Songman Kang. 2016. "Birthdays, Schooling, and Crime: Regression-Discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation." *American Economic Journal: Applied Economics*, 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. 2018. "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*, 5(9): 727–738.

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

Currie, Janet M, and W Bentley MacLeod. 2020. "Understanding Doctor Decision Making: The Case of Depression Treatment." *Econometrica*, 88(3): 847–878.

Currie, Janet, Mark Stabile, and Lauren Jones. 2014. "Do Stimulant Medications Improve Educational and Behavioral Outcomes for Children with ADHD?" *Journal of Health Economics*, 37: 58–69.

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

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

Dalsgaard, Søren, Maria Knoth Humlum, Helena Skyt Nielsen, and Marianne Simonsen. 2012. "Relative Standards in ADHD Diagnoses: The Role of Specialist Behavior." *Economics Letters*, 117(3): 663–665.

Danielson, Melissa L, Rebecca H Bitsko, Reem M Ghandour, Joseph R Holbrook, Michael D Kogan, and Stephen J Blumberg. 2018. "Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment among U.S. Children and Adolescents, 2016." *Journal of Clinical Child & Adolescent Psychology*, 47(2): 199–212.

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

Daysal, N Meltem, Marianne Simonsen, Mircea Trandafir, and Sanni Breining. 2022. "Spillover Effects of Early-Life Medical Interventions." *Review of Economics and Statistics*, 104(1): 1–16.

Depew, Briggs, and Ozkan Eren. 2016. "Born on the Wrong Day? School Entry Age and Juvenile Crime." *Journal of Urban Economics*, 96: 73–90.

Deshpande, Manasi, and Michael Mueller-Smith. 2022. "Does Welfare Prevent Crime? The Criminal Justice Outcomes of Youth Removed from SSI." *The Quarterly Journal of Economics*, 137(4): 2263–2307.

Dhuey, Elizabeth, David Figlio, Krzysztof Karbownik, and Jeffrey Roth. 2019. "School Starting Age and Cognitive Development." *Journal of Policy Analysis and Management*, 38(3): 538–578.

Dickert-Conlin, Stacy, and Amitabh Chandra. 1999. "Taxes and the Timing of Births." *Journal of political Economy*, 107(1): 161–177.

Dickert-Conlin, Stacy, and Todd Elder. 2010. "Suburban Legend: School Cutoff Dates and the Timing of Births." *Economics of Education Review*, 29(5): 826–841.

Dustan, Andrew. 2018. "Family Networks and School Choice." *Journal of Development Economics*, 134: 372–391.

Einav, Liran, Amy Finkelstein, Tamar Oostrom, Abigail Ostriker, and Heidi Williams. 2020. "Screening and Selection: The Case of Mammograms." *American Economic Review*, 110(12): 3836–70.

Elder, Todd E. 2010. "The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates." *Journal of Health Economics*, 29(5): 641–656.

Elder, Todd E, and Darren H Lubotsky. 2009. "Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers." *Journal of Human Resources*, 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. 2019. "Final Results of the Prospective FH02 Mammographic Surveillance Study of Women Aged 35–39 at Increased Familial Risk of Breast Cancer." *EClinicalMedicine*, 7: 39–46.

Evans, William N, Melinda S Morrill, and Stephen T Parente. 2010. "Measuring Inappropriate Medical Diagnosis and Treatment in Survey Data: The Case of ADHD Among School-Age Children." *Journal of Health Economics*, 29(5): 657–673.

Faraone, Stephen V, Joseph Biederman, Wei J Chen, B Krifcher, K Keenan, C Moore, S Sprich, and MT Tsuang. 1992. "Segregation Analysis of Attention Deficit Hyperactivity Disorder." *Psychiatric Genetics*.

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

Fletcher, Jason, Nicole L Hair, and Barbara L Wolfe. 2012. "Am I My Brother's Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior." National Bureau of Economic Research Working Paper 18279.

Frank, Richard G, and Thomas G McGuire. 2000. "Economics and Mental Health." In *Handbook of Health Economics*. Vol. 1, 893–954. Elsevier.

Fredriksson, Peter, and Björn Öckert. 2014. "Life-Cycle Effects of Age at School Start." *The Economic Journal*, 124(579): 977–1004.

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

Furzer, Jill, Elizabeth Dhuey, and Audrey Laporte. 2022. "ADHD Misdiagnosis: Causes and Mitigators." *Health Economics*, 31(9): 1926–1953.

Gans, Joshua S, and Andrew Leigh. 2009. "Born on the First of July: An (Un)natural Experiment in Birth Timing." *Journal of Public Economics*, 93(1-2): 246–263.

Halldner, Linda, Annika Tillander, Cecilia Lundholm, Marcus Boman, Niklas Långström, Henrik Larsson, and Paul Lichtenstein. 2014. "Relative Immaturity and ADHD: Findings From Nationwide Registers, Parent- and Self-Reports." *Journal of Child Psychology and Psychiatry*, 55(8): 897–904.

Hinshaw, Stephen P, and Richard M Scheffler. 2014. *The ADHD Explosion: Myths, Medication, Money, and Today's Push for Performance.* Oxford University Press.

Humphreys, Kathryn L, Timothy Eng, and Steve S Lee. 2013. "Stimulant Medication and Substance Use Outcomes: A Meta-Analysis." *JAMA Psychiatry*, 70(7): 740–749.

Hurwitz, Michael, Jonathan Smith, and Jessica S Howell. 2015. "Student Age and the Collegiate Pathway." *Journal of Policy Analysis and Management*, 34(1): 59–84.

Imbens, Guido W, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615–635.

Jensen, Peter S. 1999. "A 14-Month Randomized Clinical Trial of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder." *Archives of General Psychiatry*, 56(12): 1073–1086.

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

Karbownik, Krzysztof, and Umut Özek. 2023. "Setting a Good Example?: Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design." *Journal of Human Resources*, 58(5): 1567–1607.

Kawaguchi, Daiji. 2011. "Actual Age at School Entry, Educational Outcomes, and Earnings." *Journal of the Japanese and International Economies*, 25(2): 64–80.

Koehly, Laura M, June A Peters, Regina Kenen, Lindsey M Hoskins, Anne L Ersig, Natalia R Kuhn, Jennifer T Loud, and Mark H Greene. 2009. "Characteristics of Health Information Gatherers, Disseminators, and Blockers within Families at Risk of Hereditary Cancer: Implications for Family Health Communication Interventions." *American Journal of Public Health*, 99(12): 2203–2209.

Krabbe, EE, ED Thoutenhoofd, M Conradi, SJ Pijl, and L Batstra. 2014. "Birth Month as Predictor of ADHD Medication Use in Dutch School Classes." *European Journal of Special Needs Education*, 29(4): 571–578.

LaLumia, Sara, James M Sallee, and Nicholas Turner. 2015. "New Evidence on Taxes and the Timing of Birth." *American Economic Journal: Economic Policy*, 7(2): 258–293.

Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen. 2017. "School Starting Age and the Crime-Age Profile." *The Economic Journal*, 127(602): 1096–1118.

Landersø, Rasmus Kløve, Helena Skyt Nielsen, and Marianne Simonsen. 2019. "Effects of School Starting Age on the Family." *Journal of Human Resources*, 1117–9174R1.

Layton, Timothy J, Michael L Barnett, Tanner R Hicks, and Anupam B Jena. 2018. "Attention Deficit–Hyperactivity Disorder and Month of School Enrollment." *New England Journal of Medicine*, 379(22): 2122–2130.

Lee, David S, and David Card. 2008. "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics*, 142(2): 655–674.

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

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

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

McEwan, Patrick J, and Joseph S Shapiro. 2008. "The Benefits of Delayed Primary School Enrollment Discontinuity Estimates Using Exact Birth Dates." *Journal of Human Resources*, 43(1): 1–29.

Mendes, Álvaro, Milena Paneque, Liliana Sousa, Angus Clarke, and Jorge Sequeiros. 2016. "How Communication of Genetic Information within the Family Is Addressed in Genetic Counselling: A Systematic Review of Research Evidence." *European Journal of Human Genetics*, 24(3): 315–325.

Miller, Meghan, Erica D. Musser, Gregory S. Young, Brent Olson, Robert D. Steiner, and Joel T. Nigg. 2019. "Sibling Recurrence Risk and Cross-aggregation of Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder." *JAMA Pediatrics*, 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. 2013. "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*, 52(3): 250–263.

Montgomery, Susan V, Andrea M Barsevick, Brian L Egleston, Ruth Bingler, Karen Ruth, Suzanne M Miller, John Malick, Terrence P Cescon, and Mary B Daly. 2013. "Preparing Individuals to Communicate Genetic Test Results to Their Relatives: Report of a Randomized Control Trial." *Familial Cancer*, 12: 537–546.

Morrow, Richard L, E Jane Garland, James M Wright, Malcolm Maclure, Suzanne Taylor, and Colin R Dormuth. 2012. "Influence of Relative Age on Diagnosis and Treatment of Attention-Deficit/Hyperactivity Disorder in Children." *CMAJ*, 184(7): 755–762.

Moses, Tally. 2010. "Being Treated Differently: Stigma Experiences with Family, Peers, and School Staff among Adolescents with Mental Health Disorders." *Social Science & Medicine*, 70(7): 985–993.

Mullainathan, Sendhil, and Ziad Obermeyer. 2017. "Does Machine Learning Automate Moral Hazard and Error?" *American Economic Review*, 107(5): 476–80.

Nicoletti, Cheti, and Birgitta Rabe. 2019. "Sibling Spillover Effects in School Achievement." *Journal of Applied Econometrics*, 34(4): 482–501.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science*, 366(6464): 447–453.

Ohan, Jeneva L, Troy AW Visser, Melanie C Strain, and Linda Allen. 2011. "Teachers' and Education Students' Perceptions of and Reactions to Children with and without the Diagnostic Label "ADHD"." *Journal of School Psychology*, 49(1): 81–105.

Ong, Mei-Sing, and Kenneth D Mandl. 2015. "National Expenditure for False-Positive Mammograms and Breast Cancer Overdiagnoses Estimated at $4 Billion a Year." *Health Affairs*, 34(4): 576–583.

Parman, John. 2015. "Childhood Health and Sibling Outcomes: Nurture Reinforcing Nature during the 1918 Influenza Pandemic." *Explorations in Economic History*, 58(C): 22–43.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt. 2019. "Poorly Measured Confounders Are More Useful on the Left than on the Right." *Journal of Business & Economic Statistics*, 37(2): 205–216.

Pottegård, Anton, Jesper Hallas, and Helga Zoëga. 2014. "Children's Relative Age in Class and Use of Medication for ADHD: A Danish Nationwide Study." *Journal of Child Psychology and Psychiatry*, 55(11): 1244–1250.

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

Qureshi, Javaeria A. 2018*b*. "Siblings, Teachers, and Spillovers on Academic Achievement." *Journal of Human Resources*, 53(1): 272–297.

Root, Adrian, Jeremy P Brown, Harriet J Forbes, Krishnan Bhaskaran, Joseph Hayes, Liam Smeeth, and Ian J Douglas. 2019. "Association of Relative Age in the School Year With Diagnosis of Intellectual Disability, Attention-Deficit/Hyperactivity Disorder, and Depression." *JAMA pediatrics*, 173(11): 1068–1075.

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

Schulkind, Lisa, and Teny Maghakian Shapiro. 2014. "What a Difference a Day Makes: Quantifying the Effects of Birth Timing Manipulation on Infant Health." *Journal of Health Economics*, 33: 139–158.

Schwandt, Hannes, and Amelie Wuppermann. 2016. "The Youngest Get the Pill: ADHD Misdiagnosis in Germany, Its Regional Correlates and International Comparison." *Labour Economics*, 43: 72–86.

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

Socialstyrelsen. 2012. "Användning av centralstimulantia vid adhd." *Socialstyrelsens Rapporter.*

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

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

Socialstyrelsen. 2022. "Nationella riktlinjer för vård och stöd vid adhd och autism." *Socialstyrelsen.*

Tarver, Joanne, David Daley, and Kapil Sayal. 2014. "Attention-Deficit Hyperactivity Disorder (ADHD): An Updated Review of the Essential Facts." *Child: Care, Health and Development*, 40(6): 762–774.

Thapar, Anita, and Miriam Cooper. 2016. "Attention Deficit Hyperactivity Disorder." *The Lancet*, 387(10024): 1240–50.

Thomas, Rae, Sharon Sanders, Jenny Doust, Elaine Beller, and Paul Glasziou. 2015. "Prevalence of Attention-Deficit/Hyperactivity Disorder: A Systematic Review and Meta-Analysis." *Pediatrics*, 135(4): e994–e1001.

Visser, Susanna N, Melissa L Danielson, Rebecca H Bitsko, Joseph R Holbrook, Michael D Kogan, Reem M Ghandour, Ruth Perou, and Stephen J Blumberg. 2014. "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*, 53(1): 34–46.

Whitely, Martin, Melissa Raven, Sami Timimi, Jon Jureidini, John Phillimore, Jonathan Leo, Joanna Moncrieff, and Patrick Landman. 2018. "Attention Deficit Hyperactivity Disorder Late Birthdate Effect Common in Both High and Low Prescribing International Jurisdictions: Systematic Review." *Journal of Child Psychology and Psychiatry*.

Wilens, Timothy E, Stephen V Faraone, Joseph Biederman, and Samantha Gunawardene. 2003. "Does Stimulant Therapy of Attention-Deficit/Hyperactivity Disorder Beget Later Substance Abuse? A Meta-Analytic Review of the Literature." *Pediatrics*, 111(1): 179–185.

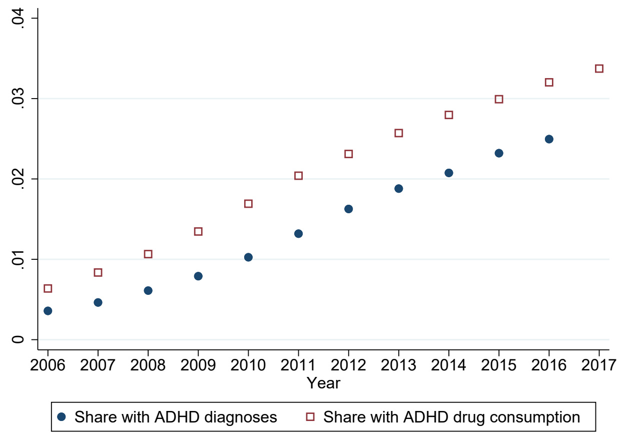
Wilson, Brenda J, Karen Forrest, Edwin R van Teijlingen, Lorna McKee, Neva Haites, Eric Matthews, and Sheila A Simpson. 2004. "Family Communication about Genetic Risk: The Little That Is Known." *Public Health Genomics*, 7(1): 15–24.

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

Zoëga, Helga, Unnur A Valdimarsdóttir, and Sonia Hernández-Díaz. 2012. "Age, Academic Performance, and Stimulant Prescribing for ADHD: A Nationwide Cohort Study." *Pediatrics*, 130(6): 1012–1018.

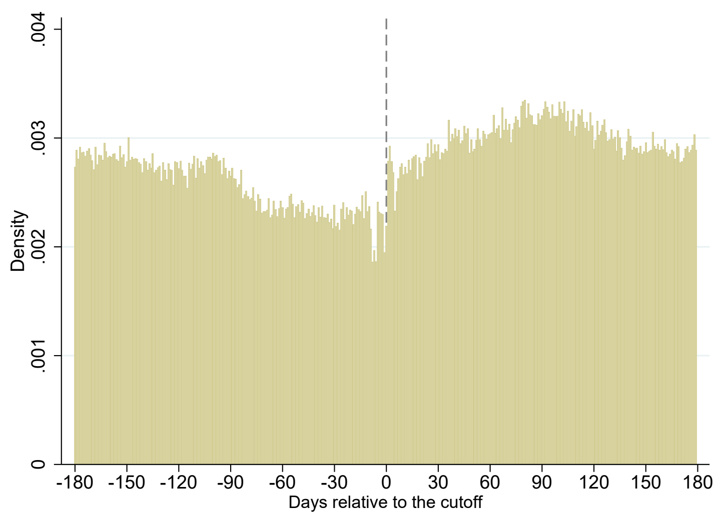
7 Figures and Tables

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



Note: The sample includes children between the ages of 6 and 19 who are born in Sweden. 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.

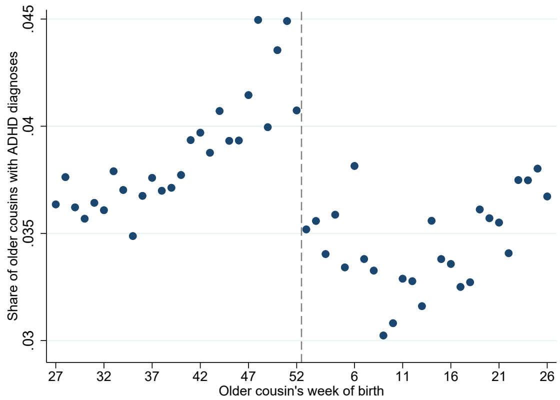
Figure 2: Distribution of Births at the Daily Level, Older Cousins



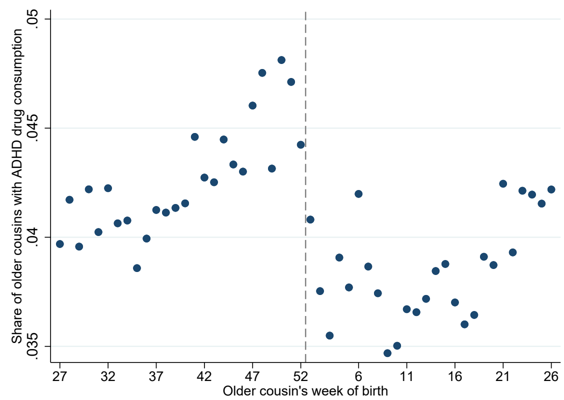
Note: The sample of analysis is the universe of cousin pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. The figure shows a histogram of the distribution of older cousins' births at the daily level, with a bandwidth of 180 days around the cutoff (January 1).

Figure 3: ADHD Diagnoses and Drug Treatment by Own Week of Birth, Older Cousins Only

(a) Diagnoses



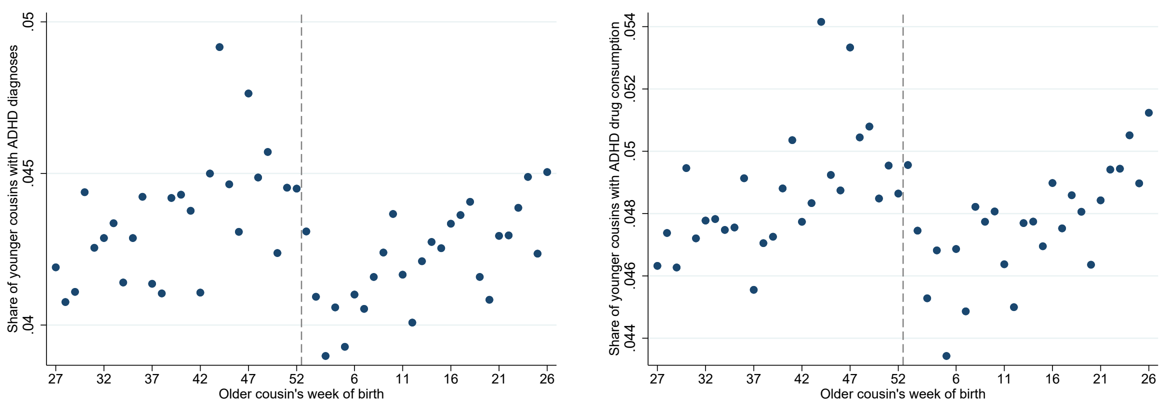
(b) Drug Treatment



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

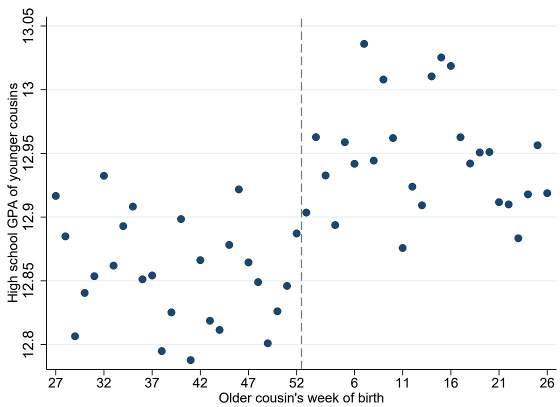
(a) Diagnoses (b) Drug Treatment



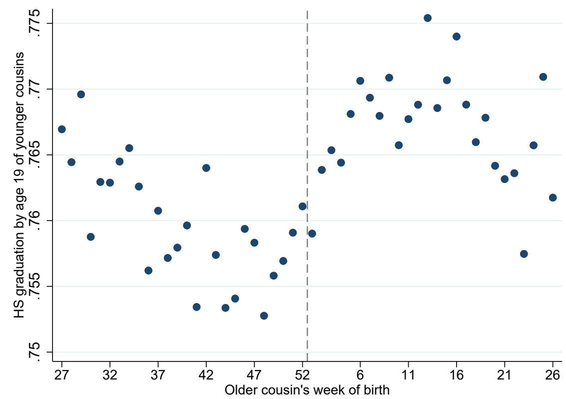
Note: The sample of analysis is the universe of cousin pairs born in Sweden, where the older cousin is born between July 1985 and June 1996. These graphs plot ADHD-related outcomes for younger cousins (on the y-axes) by the birth week of the older cousin (on the x-axes). Sub-figure (a) plots the share 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

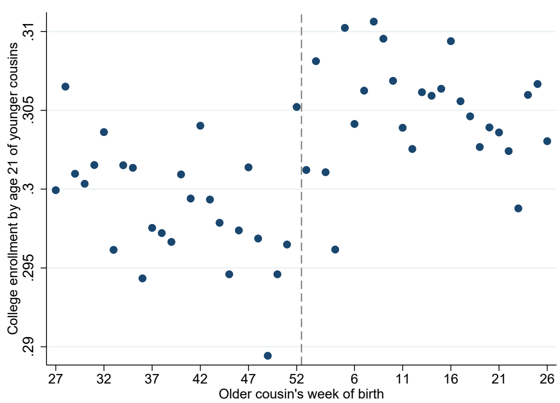
(a) HS GPA



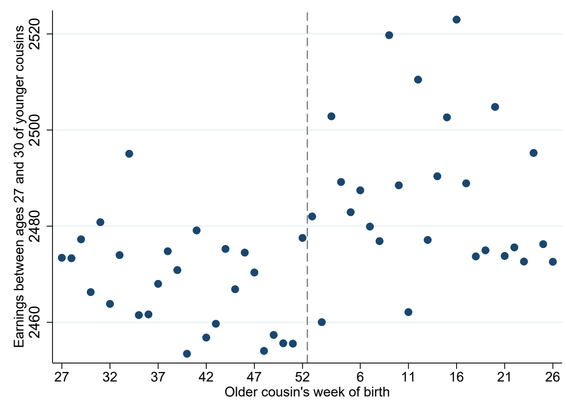
(b) HS Graduation on Time



(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 3 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	0.0071***	0.0072***
[Own Relative Age Effect]	(0.0019)	(0.0017)
Mean(Y)	0.037	0.041
N	221,660	221,660
Panel B: Full Covariates		
OC born before the cutoff	0.0065***	0.0066***
[Own Relative Age Effect]	(0.0018)	(0.0017)
Mean(Y)	0.037	0.041
N	221,660	221,660

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	0.0042***	0.0030**
[Spillover Effect]	(0.0012)	(0.0013)
YC born before the cutoff	0.0142***	0.0158***
[Own Relative Age Effect]	(0.0012)	(0.0014)
Mean(Y)	0.043	0.048
N	432,903	432,903
Panel B: Full Covariates		
OC born before the cutoff	0.0039***	0.0028**
[Spillover Effect]	(0.0012)	(0.0013)
YC born before the cutoff	0.0119***	0.0131***
[Own Relative Age Effect]	(0.0013)	(0.0014)
Mean(Y)	0.043	0.048
N	432,903	432,903

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)	(0.0003)
Mean(Y)	0.003	0.001	0.001
N	432,903	432,903	432,903
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	0.0007*	0.0001	0.0005**
[Own Relative Age Effect]	(0.0004)	(0.0002)	(0.0003)
Mean(Y)	0.003	0.001	0.001
N	432,903	432,903	432,903

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	-0.0869**	-0.0060**	-0.0068*	-21.7847
[Spillover Effect]	(0.0369)	(0.0030)	(0.0037)	(14.7029)
YC born before the cutoff	-0.4695***	-0.0729***	-0.0265***	-132.9467***
[Own Relative Age Effect]	(0.0404)	(0.0029)	(0.0031)	(13.9655)
Mean(Y)	12.909	0.763	0.302	2476.607
N	346,794	402,243	401,254	217,673
Panel B: Full Covariates				
OC born before the cutoff	-0.0637**	-0.0044	-0.0040	-20.6018
[Spillover Effect]	(0.0314)	(0.0029)	(0.0032)	(14.0851)
YC born before the cutoff	-0.5129***	-0.0699***	-0.0231***	-125.0316***
[Own Relative Age Effect]	(0.0371)	(0.0029)	(0.0031)	(13.1117)
Mean(Y)	12.909	0.763	0.302	2476.607
N	346,794	402,243	401,254	217,673

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.[41] 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.

[41] 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 The Healthcare Costs of Spillover Diagnoses

In this section, we calculate the total healthcare costs associated with the spillover ADHD diagnoses. This is the sum of the spillover-induced ADHD treatment costs borne by the public insurer and the spillover-induced private out-of-pocket costs of ADHD treatment.

We follow Deshpande and Mueller-Smith (2022) in using "intensive margin" outcomes when calculating these spillover costs. Appendix Table C17 presents results for impacts on several intensive margin measures of ADHD treatment as outcomes among the younger cousins. We find a spillover effect on the total number of unique ADHD drugs used by younger cousins in the first year following diagnosis of 0.0023 (column (1)).[42] Further, we find spillover effects on the number of mental health-related outpatient visits in the first, second, and third years following the diagnosis of 0.012, 0.0087, and 0.0064, respectively (columns (2)-(4)).

Using an estimate of the total cost of a typical ADHD drug in Sweden of $852 (in 2019 USD, calculated below), coupled with our sample size of 616,242 unique younger cousins, yields a total cost of the additional drug spending in the first year induced by spillover diagnoses of $0.0023 \times 852 \times 616242$, or $1.208 million.

Using an estimate of the total cost of one psychotherapy visit in Sweden of $153 (in 2019 USD, calculated below), coupled with our sample size of 616,242 unique younger cousins, implies a total cost of the additional mental health-related outpatient visits in the first three years induced by spillover diagnoses of $(0.012 + 0.0087 + 0.0064) \times 153 \times 616242$, or $2.555 million.

This yields a total cost of $1.208+$2.555=$3.763 million. This cost stems from $2,403$ spillover diagnoses,[43] yielding a total healthcare cost per spillover diagnosis of **$1,565 (in 2019 USD)**. Of this, $502 is the implied cost of spillover-induced drug consumption in the first year post-diagnosis ($1.208 million divided by the sample size of $616,242$), and $1,063 is the implied cost of spillover-induced psychotherapy over the first three years post-diagnosis ($2.555 million divided by the sample size of $616,242$).

[42] Appendix Table C16 documents a spillover on extensive margin drug consumption in the first year following diagnosis, but no statistically significant spillover on extensive margin drug consumption in subsequent two years. Thus, in Appendix Table C17 we restrict attention to the intensive margin effect on drug consumption in the first year following the diagnosis.

[43] The spillover estimate on the number of diagnoses, reported in Table 2, is 0.0039, which we multiply by the number of younger cousins, $616,242$, to obtain the total number of spillover diagnoses.

We can express this number as a share of the total cost of treatment for an average ADHD patient. The average annual cost of drug treatment, among all patients taking ADHD medication in Sweden, is $1,078 (see calculations below). Thus, a spillover diagnosis induces drug spending of approximately one-half of a typical ADHD patient in the first year following diagnosis ($\frac{502}{1{,}078} = 0.47$). In the second and third year following diagnosis, the average ADHD drug user's adherence shares (reported in Appendix Table C11) are 0.755 in the second year and 0.594 in the third year. Thus, for the average patient taking ADHD drugs, the total cost of drug treatment in the first three years is given by $(1+0.755+0.594) \times 1,078 = \$2,522$. This means that a spillover diagnosis induces drug spending of approximately one-fifth of a typical ADHD patient in the first three years following diagnosis ($\frac{502}{2{,}522} = 0.20$). We note that these shares represent only drug spending and not spending on psychotherapy. Because spillover diagnoses are associated with significant drug spending only in the first year (and not in the second and third), whereas spillover diagnoses are associated with psychotherapy costs over years two and three as well, calculating the share of the total cost that is associated with a spillover diagnosis from the drug costs alone yields a conservative estimate of the total share of healthcare costs associated with spillover diagnoses.

The total cost of ADHD treatment We calculate the total cost of ADHD treatment in 2019, a year close to the end of our sample period for which we have information about treatment costs in Sweden.

<u>Cost of ADHD drug treatment:</u> The total costs borne by the public insurer (regions) for all ADHD drugs consumed in Sweden in 2019, amounted to 1.026 billion SEK.[44] In the same year, 130,000 individuals received at least one ADHD drug,[45] which yields an average insurer drug cost of $7,892$ SEK per treated patient.

The private out-of-pocket drug cost is $2,300$ SEK per patient (assuming that the patient reaches the maximum out-of-pocket expense in 2019). This yields an average total cost of $7,892 + 2,300 = 10,192$ SEK, or $\$1,078$ USD,[46] per patient treated with ADHD drugs.

[44]See `https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2022-4-7858.pdf`, accessed in February 2024.

[45]See `https://www.lakemedelsvarlden.se/okad-forskrivning-av-adhd-lakemedel/`, accessed in February 2024.

[46]This conversion uses the average SEK USD exchange rate in 2019. See `https://www.exchangerates.org.uk/USD-SEK-spot-exchange-rates-history-2019.html`, accessed in February 2024.

Some patients are treated with only one ADHD drug (one ATC code) whereas other patients take more than one unique ADHD drug in a year. The typical patient under the age of 25 who takes ADHD medication takes 1.265 unique drugs per year.[47] This gives an annual total cost per unique ADHD drug of $\frac{1{,}078}{1.265} = \$852$ USD.

Cost of psychotherapy: The public insurer's cost for one pediatric psychotherapy visit is $1,350$ SEK.[48] Assuming a (typical) private co-pay of 100 SEK yields a total cost of $1,450$ SEK, or \$153 USD (using the same exchange rate as above).

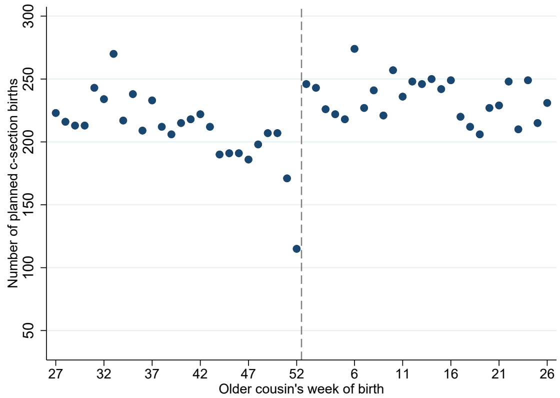
[47] See `https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2018-3-30.pdf`, accessed in February 2024, for the share of patients consuming various ADHD drugs, by gender, in the age ranges 5-9, 10-17, and 18-24, respectively. We take the average number of unique ADHD drugs across the six groups, weighted by the number of patients consuming ADHD drugs in each group. Note that this number is for 2017, whereas the cost estimates are from 2019; thus we implicitly assume that the average number of drugs consumed remains constant from 2017 to 2019.

[48] See `https://sodrasjukvardsregionen.se/download/regionala-priser-och-ersattningar-for-sodra-sjukvardsregionen-2019/?wpdmdl=10574&refresh=65cd7d3a023b71707965754`, accessed in February 2024. The estimate represents the cost in the Southern Hospital Region, which includes the regions of Skåne, Blekinge, Kronoborg, and Halland.

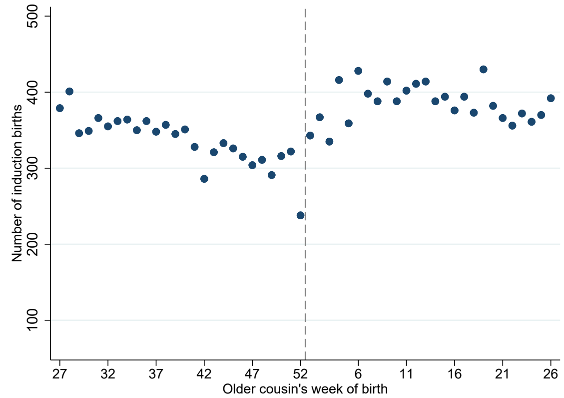
C Additional Results

Figure C1: Number of Births By Week Among Older Cousins Sample: Planned C-Sections, Inductions, and Non-Induced Vaginal Deliveries

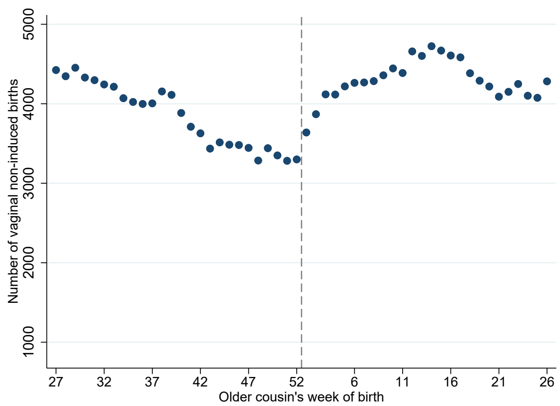
(a) Planned C-Sections



(b) Inductions



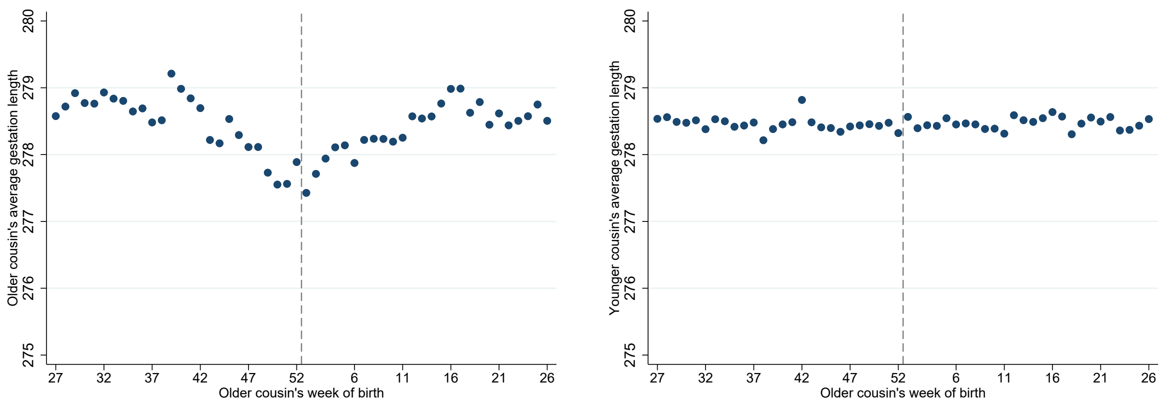
(c) Non-Induced Vaginal Deliveries



Note: See notes under Figure 3 for more information about the sample of older cousins. Sub-figure (a) plots the number of births by week of the year that were planned cesarian (c-section) deliveries. Sub-figure (b) plots the number of births by week of the year that were induced. Sub-figure (c) plots the number of births by week of the year that were non-induced vaginal deliveries.

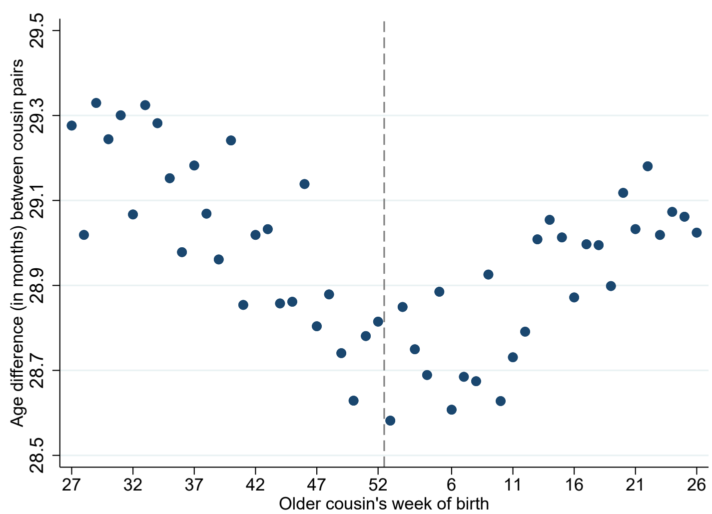
Figure C2: Average Gestation Length by Older Cousin's Week of Birth

(a) Gestation Length, Older Cousins (b) Gestation Length, Younger Cousins



Note: See notes under Figure 3 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.

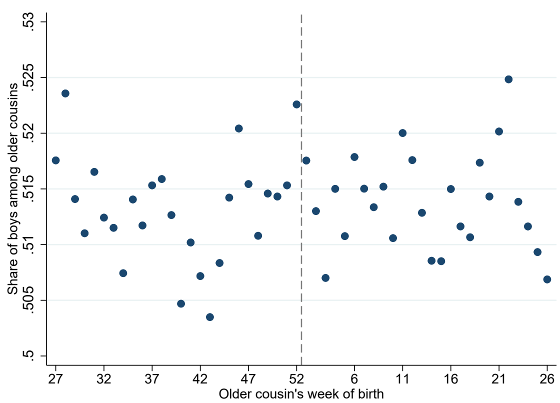
Figure C3: Age Difference Between Cousins by Week of Birth of Older Cousin



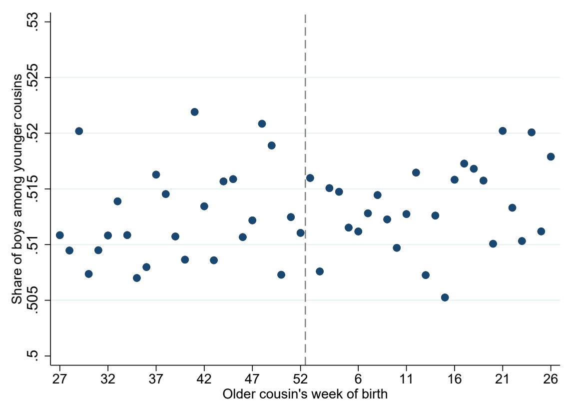
Note: See notes under Figure 3 for more information about the sample. The figure plots the average age difference between cousins (in months) by the birth week of the older cousin.

Figure C4: Cousin Gender Composition by Week of Birth of Older Cousin

(a) Share Boys, Older Cousins



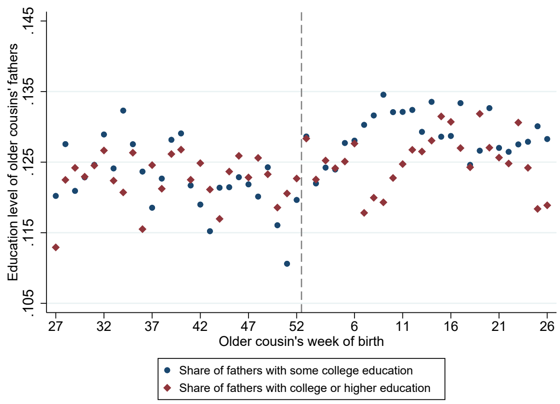
(b) Share Boys, Younger Cousins



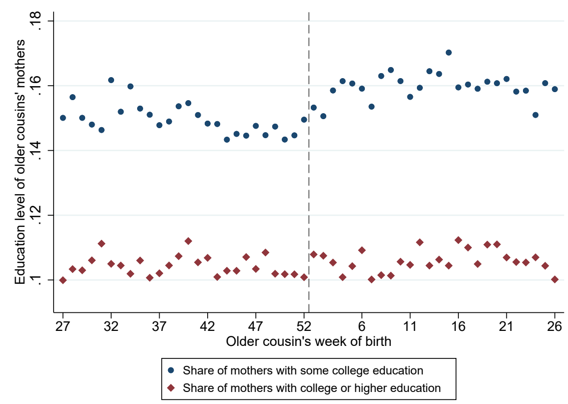
Note: See notes under Figure 3 for more information about the sample. These figures plot the share of boys among older and younger cousins by the birth week of the older cousin. Sub-figure (a) plots the share of boys among older cousins in our main sample of cousin pairs. Sub-figure (b) plots the share of boys among younger cousins in our main sample of cousin pairs.

Figure C5: Older and Younger Cousins' Parental Education Level by Own Week of Birth

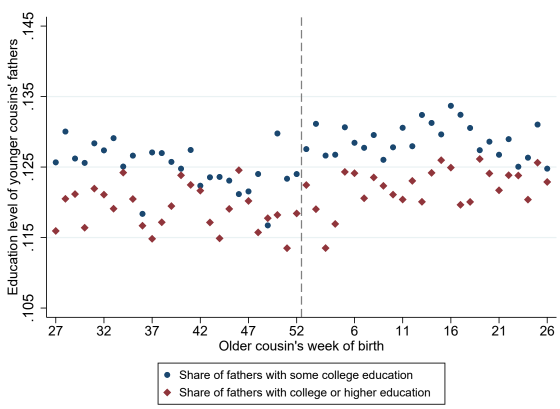
(a) Fathers, Older Cousins



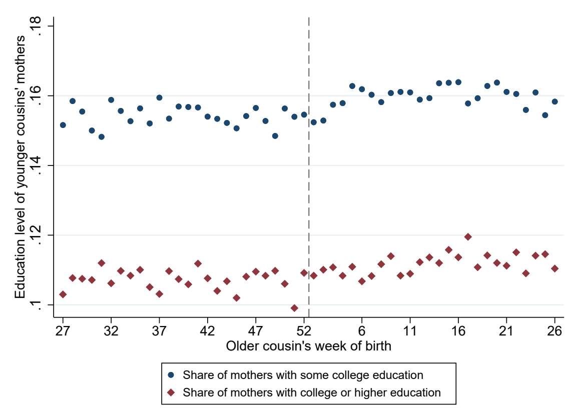
(b) Mothers, Older Cousins



(c) Fathers, Younger Cousins



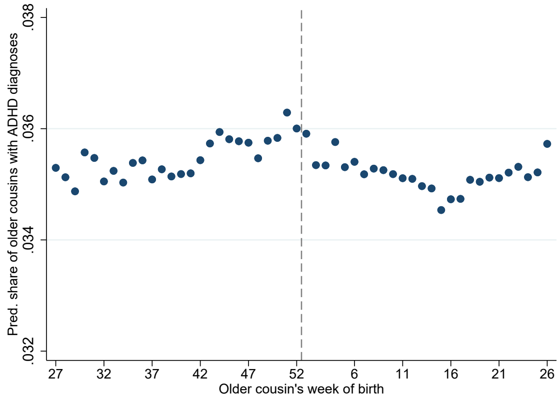
(d) Mothers, Younger Cousins



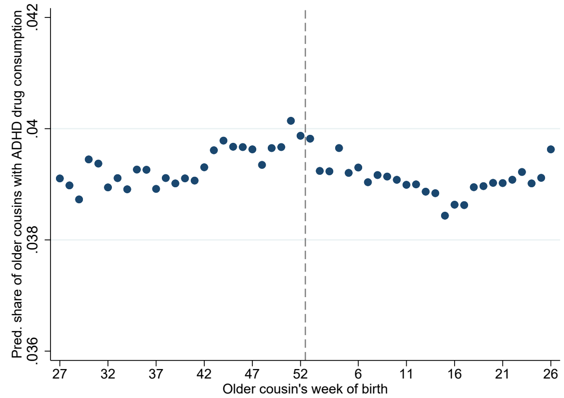
Note: See notes under Figure 3 for more information about the sample. Sub-figure (a) plots the share of older cousins' fathers with some college education (in blue) and the share of older cousins' fathers with college or higher level of education (in red), by the older cousin's week of birth. Sub-figure (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 C6: Predicted ADHD Outcomes by Week of Birth of Older Cousin

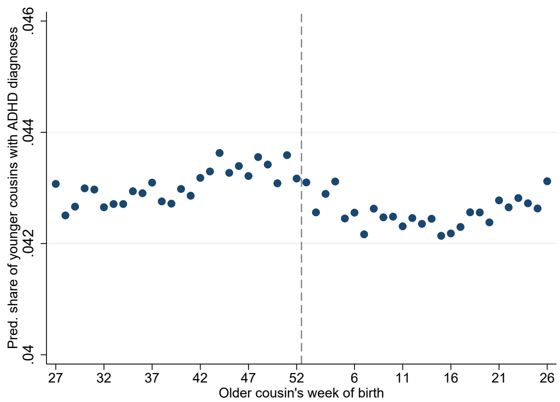
(a) Pred. Diagnoses, Older Cousins



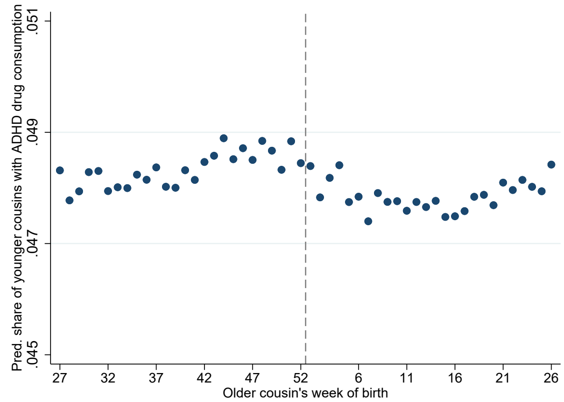
(b) Pred. Drug Treatment, Older Cousins



(c) Pred. Diagnoses, Younger Cousins

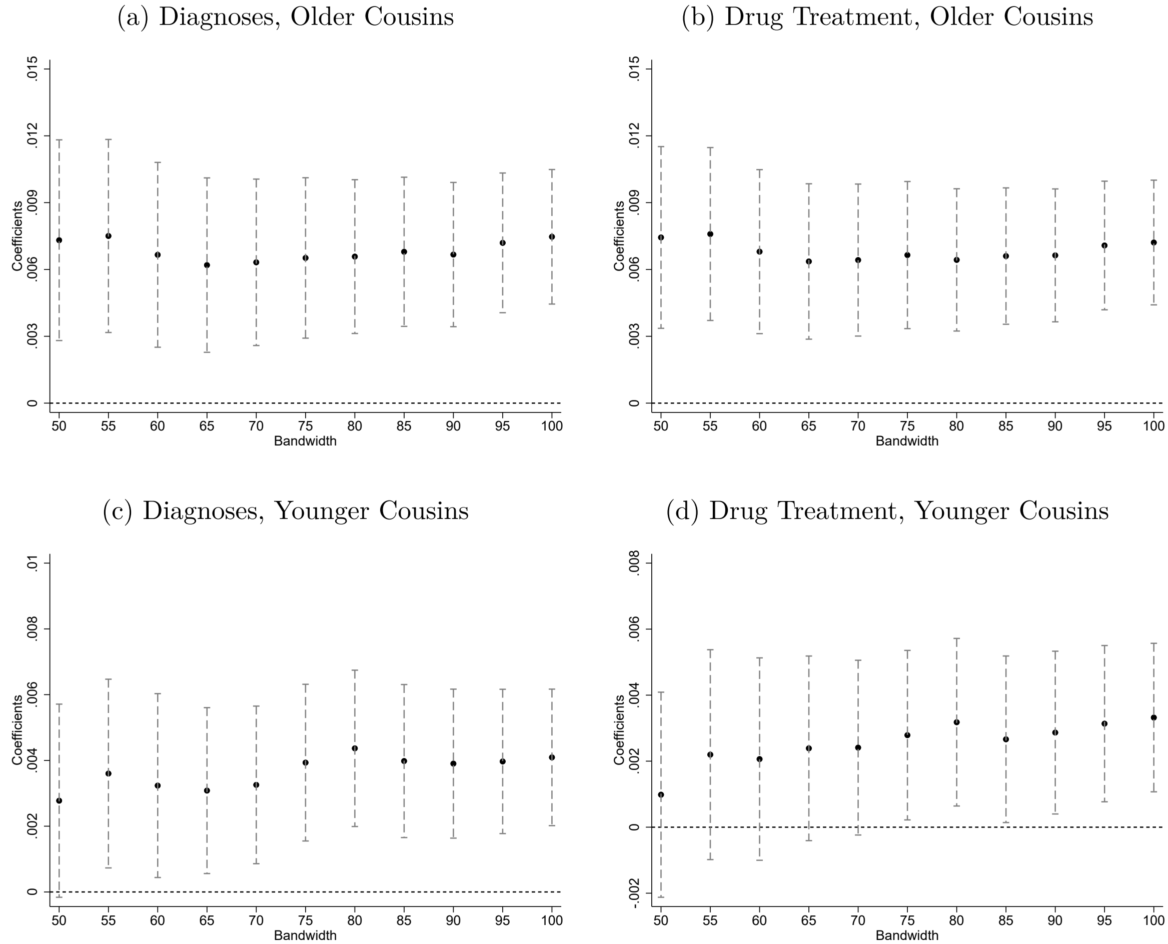


(d) Pred. Drug Treatment, Younger Cousins



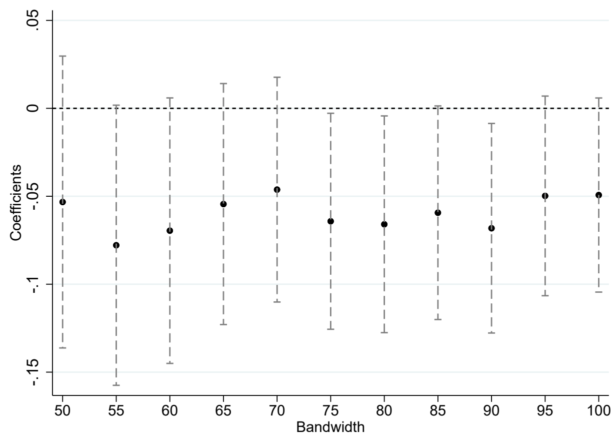
Note: See notes under Figure 3 for more information about the sample. These graphs plot predicted ADHD-related outcomes for the older cousins and younger cousins, respectively, by the birth week of the older cousin. The predicted outcomes of older and younger cousins are constructed by regressing each 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 C7: Effect of Older Cousin Being Born Before Cutoff on Own and Younger Cousin's ADHD Diagnoses and Drug Treatment, with Varying Bandwidth



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.

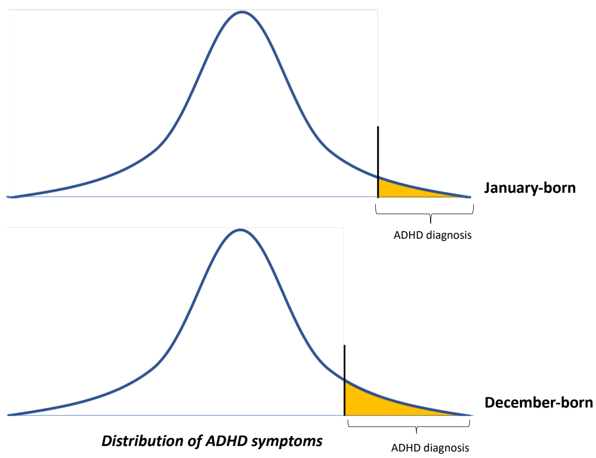
Figure C8: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's High School GPA, with Varying Bandwidth



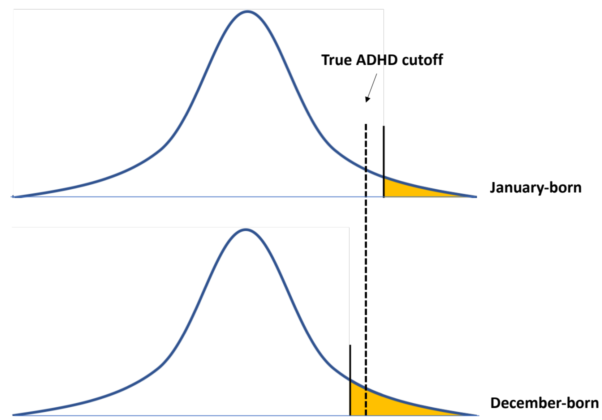
Note: This figure plots regression coefficients and 95% confidence intervals from estimating model (2), varying the bandwidth from 50 days to 100 days. See notes under Table 2 for more details about the sample, specifications, and control variables. The outcome variable is the younger cousin's high school GPA.

Figure C9: Stylized Framework for Interpreting ADHD Gap; Three Possible Interpretations

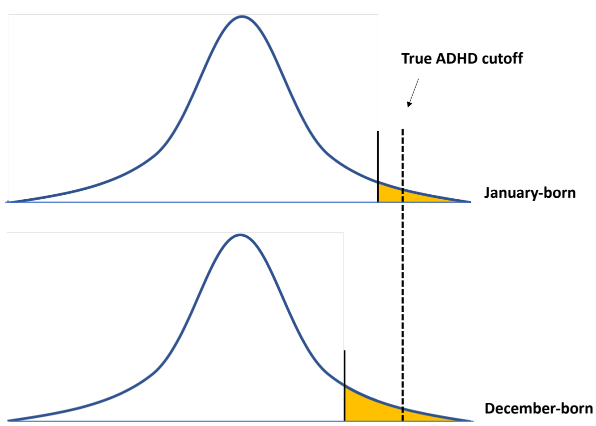
(a) Baseline Framework



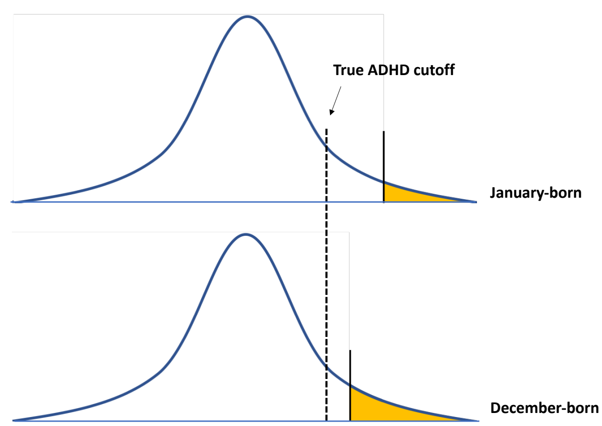
(b) Over-Diagnosis in Dec., Under-Diagnosis in Jan.



(c) More Over-Diagnosis in December than January



(d) 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 of sub-figures (b), (c), and (d) 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 C1: Sample Means of Key Variables

Panel A: Older Cousins	Full Sample	Jul-Dec	Jan-Jun
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	575,213	267,996	307,217
Panel B: Younger Cousins	Full Sample	Jul-Dec	Jan-Jun
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	29.051	28.889
Observations	1,122,747	524,813	597,934

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 in July-December and January-June, respectively.

Table C2: 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	-0.0830	0.0077*	-0.0014	-0.0019	
	(0.1250)	(0.0044)	(0.0026)	(0.0021)	
Mean(Y)	278.044	0.514	0.123	0.104	
N	221,660	221,660	221,660	221,660	
Panel B: Younger Cousins					
OC born before the cutoff	-0.1483	-0.0001	0.0004	-0.0002	0.1446*
	(0.0982)	(0.0029)	(0.0024)	(0.0021)	(0.0871)
Mean(Y)	278.440	0.513	0.120	0.108	28.788
N	432,903	432,903	432,903	432,903	432,903

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 C3: Effect of Older Cousin Being Born Before Cutoff on Own Mother's Labor Market Outcomes and Marital Status

	(1) Employed	(2) Work Income	(3) Married
Panel A: No Covariates			
OC born before the cutoff	-0.0101***	-24.5106**	-0.0251***
[Own Relative Age Effect]	(0.0024)	(10.1990)	(0.0034)
Mean(Y)	0.892	1398.085	0.623
N	221,660	220,541	220,465
Panel B: Full Covariates			
OC born before the cutoff	-0.0088***	-16.3277*	-0.0219***
[Own Relative Age Effect]	(0.0022)	(8.7692)	(0.0034)
Mean(Y)	0.892	1398.085	0.623
N	221,660	220,541	220,465

Notes: Each column reports results from separate regressions estimating model (1). See notes under Table 1 for more information on the analysis sample, specifications, and control variables. The dependent variable in column (1) is an indicator equal to one if the older cousin's mother is employed when the older cousin is 7 years old. The dependent variable in column (2) is the older cousin's mother's work income when the older cousin is 7 years old. The dependent variable in column (3) is an indicator equal to one if the older cousin's mother is married when the older cousin is 7 years old. 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 C4: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's Mother's Labor Market Outcomes and Marital Status

	(1) Employed	(2) Work Income	(3) Married
Panel A: No Covariates			
OC born before the cutoff	-0.0027	-6.3035	-0.0045
[Spillover Effect]	(0.0028)	(8.3328)	(0.0045)
YC born before the cutoff	-0.0131***	-57.6000***	-0.0245***
[Own Relative Age Effect]	(0.0023)	(7.4991)	(0.0032)
Mean(Y)	0.875	1229.503	0.607
N	432,903	431,286	430,396
Panel B: Full Covariates			
OC born before the cutoff	-0.0020	-3.2626	-0.0020
[Spillover Effect]	(0.0025)	(6.9567)	(0.0043)
YC born before the cutoff	-0.0056**	-20.3772***	-0.0163***
[Own Relative Age Effect]	(0.0022)	(6.1279)	(0.0031)
Mean(Y)	0.875	1229.503	0.607
N	432,903	431,286	430,396

Notes: Each column reports results from a separate regression estimating model (2). See notes under Table 2 for more information on the analysis sample, specifications, and control variables. The dependent variable in column (1) is an indicator equal to one if the younger cousin's mother is employed when the older cousin is 7 years old. The dependent variable in column (2) is the younger cousin's mother's work income when the older cousin is 7 years old. The dependent variable in column (3) is an indicator equal to one if the younger cousin's mother is married when the older cousin is 7 years old. 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 C5: 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	0.0088***	0.0071***	0.0063***	0.0086***	0.0069***	0.0064***
[Own Relative Age Effect]	(0.0011)	(0.0017)	(0.0024)	(0.0011)	(0.0016)	(0.0021)
Mean(Y)	0.036	0.036	0.036	0.040	0.040	0.040
N	575,213	575,213	575,213	575,213	575,213	575,213
Panel B: Younger Cousins						
OC born before the cutoff	0.0041***	0.0038***	0.0031**	0.0038***	0.0027**	0.0013
[Spillover Effect]	(0.0008)	(0.0012)	(0.0015)	(0.0009)	(0.0013)	(0.0016)
YC born before the cutoff	0.0125***	0.0119***	0.0126***	0.0134***	0.0120***	0.0128***
[Own Relative Age Effect]	(0.0008)	(0.0013)	(0.0017)	(0.0009)	(0.0013)	(0.0017)
Mean(Y)	0.043	0.043	0.043	0.048	0.048	0.048
N	1,122,747	1,122,747	1,122,747	1,122,747	1,122,747	1,122,747

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 C6: Effect of Older Cousin Being Born Before Cutoff on Own ADHD Diagnosis and Drug Treatment, Non-Induced Vaginal Deliveries Only

	(1) ADHD Diag	(2) ADHD Drug
Panel A: No Covariates		
OC born before the cutoff	0.0128***	0.0129***
[Own Relative Age Effect]	(0.0032)	(0.0031)
Mean(Y)	0.047	0.051
N	81,246	81,246
Panel B: Full Covariates		
OC born before the cutoff	0.0116***	0.0116***
[Own Relative Age Effect]	(0.0032)	(0.0031)
Mean(Y)	0.047	0.051
N	81,246	81,246

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 only include cousin pairs in which the older cousin was born via a non-induced vaginal delivery. 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 C7: 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	0.0063**	0.0080***
[Own Relative Age Effect]	(0.0027)	(0.0025)
Mean(Y)	0.037	0.041
N	182,972	182,972
Panel B: Younger Cousins		
OC born before the cutoff	0.0057***	0.0054**
[Spillover Effect]	(0.0021)	(0.0023)
YC born before the cutoff	0.0124***	0.0134***
[Own Relative Age Effect]	(0.0014)	(0.0016)
Mean(Y)	0.043	0.048
N	356,614	356,614

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 C8: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment, Non-Induced Vaginal Deliveries Only

	(1) ADHD Diag	(2) ADHD Drug
Panel A: No Covariates		
OC born before the cutoff	0.0087***	0.0071**
[Spillover Effect]	(0.0028)	(0.0032)
YC born before the cutoff	0.0203***	0.0247***
[Own Relative Age Effect]	(0.0028)	(0.0029)
Mean(Y)	0.052	0.059
N	125,508	125,508
Panel B: Full Covariates		
OC born before the cutoff	0.0081***	0.0065**
[Spillover Effect]	(0.0028)	(0.0031)
YC born before the cutoff	0.0182***	0.0219***
[Own Relative Age Effect]	(0.0028)	(0.0029)
Mean(Y)	0.052	0.059
N	125,508	125,508

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 only include cousin pairs in which the older cousin was born via a non-induced vaginal delivery. 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 C9: 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	0.0027	0.0026	0.0038*	0.0038*	0.0028	0.0078***	0.0073***	0.0087***	0.0087***	0.0085***
[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	0.044	0.042	0.042	0.043	0.042	0.043	0.043	0.043	0.043
N	256,122	273,822	215,296	215,296	232,966	123,700	135,532	106,976	106,976	114,755
Left BW	45.25	32.72	38.18	38.18	38.18	22.55	16.30	19.02	19.02	19.02
Right BW	45.25	68.96	38.18	38.18	45.25	22.55	34.36	19.02	19.02	22.55
Panel B: ADHD Drug Treatment										
OC born before the cutoff	0.0006	0.0015	0.0016	0.0016	0.0010	0.0057**	0.0056**	0.0062**	0.0062**	0.0061**
[Spillover Effect]	(0.0020)	(0.0019)	(0.0021)	(0.0021)	(0.0020)	(0.0028)	(0.0028)	(0.0030)	(0.0030)	(0.0029)
Mean(Y)	0.048	0.049	0.048	0.048	0.048	0.048	0.049	0.049	0.049	0.049
N	238,753	271,197	209,449	209,449	222,159	118,239	132,990	101,505	101,505	109,294
Left BW	42.46	32.76	37.66	37.66	37.66	21.16	16.32	18.77	18.77	18.77
Right BW	42.46	67.86	37.66	37.66	42.46	21.16	33.81	18.77	18.77	21.16

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

Table C10: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's High School GPA, Robustness Checks

	(1)	(2)	(3)	(4)	(5)
	Doughnut-RD	Linear	Quadratic	Cubic	Non-Induced Vaginal
OC born before the cutoff	-0.0780*	-0.0780***	-0.0475	-0.0547	0.0778
[Spillover Effect]	(0.0461)	(0.0211)	(0.0312)	(0.0429)	(0.0650)
YC born before the cutoff	-0.5531***	-0.5077***	-0.4370***	-0.3848***	-0.4074***
[Own Relative Age Effect]	(0.0412)	(0.0227)	(0.0301)	(0.0421)	(0.0565)
Mean(Y)	12.911	12.907	12.907	12.907	13.356
N	285,832	898,820	898,820	898,820	89,050

Notes: Each column reports results from a separate regression. In column (1), we use a bandwidth of 75 days and exclude all cousins pairs with older cousins born in the two-week bandwidth around the cutoff (January 1st). In columns (2), (3), and (4), we use a global bandwidth and include linear, quadratic, and cubic polynomials of the running variable, respectively. In column (5), we use a bandwidth of 75 days and only include cousin pairs in which the older cousin was born via a non-induced vaginal delivery. In all columns, we include the same set of control variables as in Panel B of Tables 1 and 2. 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 C11: ADHD Drug Adherence Rates Over Time

	Adherence mean
1st Year	1.000
2nd Year	0.755
3rd Year	0.594
4th Year	0.482
5th Year	0.395
N	15130

Notes: This table reports on ADHD drug adherence over five years following an ADHD diagnosis in our sample. The sub-sample used for these calculations includes all individuals in our baseline sample who receive an ADHD diagnosis after July 2005 and who obtain an ADHD drug within one year of diagnosis. We then report what share of these individuals also have an ADHD drug claim two, three, four, and five years later.

Table C12: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment, Heterogeneity by Gender Composition of Sibling Parents

	(1) ADHD Diag	(2) ADHD Drug
Panel A: Mom-Mom		
OC born before the cutoff	0.0045*	0.0041
[Spillover Effect]	(0.0023)	(0.0027)
YC born before the cutoff	0.0056**	0.0074***
[Own Relative Age Effect]	(0.0026)	(0.0027)
Mean(Y)	0.042	0.048
N	114,130	114,130
Panel B: Mom-Dad		
OC born before the cutoff	0.0065***	0.0048**
[Spillover Effect]	(0.0019)	(0.0020)
YC born before the cutoff	0.0142***	0.0158***
[Own Relative Age Effect]	(0.0019)	(0.0020)
Mean(Y)	0.043	0.048
N	210,202	210,202
Panel C: Dad-Dad		
OC born before the cutoff	-0.0014	-0.0023
[Spillover Effect]	(0.0026)	(0.0027)
YC born before the cutoff	0.0141***	0.0139***
[Own Relative Age Effect]	(0.0026)	(0.0027)
Mean(Y)	0.043	0.048
N	109,173	109,173

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. Panel A restricts the sample to cousin pairs in which the related parents are both mothers (i.e., sisters); Panel B restricts the sample to cousin pairs in which the related parents are a mother and a father (i.e., a sister and a brother); Panel C restricts the sample to cousin pairs in which the related parents are both fathers (i.e., brothers). 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 C13: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment, Heterogeneity by Older Cousin's Household Income

	(1) ADHD Diag	(2) ADHD Drug
Panel A: Below Median Household Income		
OC born before the cutoff	0.0049**	0.0021
[Spillover Effect]	(0.0019)	(0.0022)
YC born before the cutoff	0.0139***	0.0148***
[Own Relative Age Effect]	(0.0020)	(0.0019)
Mean(Y)	0.049	0.055
N	216,562	216,562
Panel B: Above Median Household Income		
OC born before the cutoff	0.0030*	0.0035**
[Spillover Effect]	(0.0016)	(0.0017)
YC born before the cutoff	0.0098***	0.0112***
[Own Relative Age Effect]	(0.0017)	(0.0019)
Mean(Y)	0.036	0.042
N	216,341	216,341

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. To measure household income, we take the average of household income in the year of birth, the year after birth, and the second year after birth. Panel A restricts to cousin pairs in which the older cousin's household income is below the median, while Panel B restricts to cousin pairs in which the older cousin's household income is above the median. 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 C14: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment, Heterogeneity by Older Cousin's Mother's Foreign-Born Status

	(1) ADHD Diag	(2) ADHD Drug
Panel A: Foreign-Born		
OC born before the cutoff	0.0006	-0.0013
[Spillover Effect]	(0.0045)	(0.0044)
YC born before the cutoff	0.0113***	0.0165***
[Own Relative Age Effect]	(0.0042)	(0.0043)
Mean(Y)	0.052	0.056
N	42,390	42,390
Panel B: Not Foreign-Born		
OC born before the cutoff	0.0042***	0.0032**
[Spillover Effect]	(0.0012)	(0.0013)
YC born before the cutoff	0.0119***	0.0126***
[Own Relative Age Effect]	(0.0013)	(0.0015)
Mean(Y)	0.042	0.047
N	390,513	390,513

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. Panel A restricts to cousin pairs in which the older cousin's mother is born outside of Sweden, while Panel B restricts to cousin pairs in which the older cousin's mother is Swedish-born. 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 C15: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Diagnosis and Drug Treatment, Heterogeneity by Older Cousin's Mother's Education

	(1) ADHD Diag	(2) ADHD Drug
Panel A: No College		
OC born before the cutoff	0.0044***	0.0036**
[Spillover Effect]	(0.0014)	(0.0016)
YC born before the cutoff	0.0118***	0.0134***
[Own Relative Age Effect]	(0.0016)	(0.0017)
Mean(Y)	0.047	0.052
N	326,113	326,113
Panel B: Some or Full College		
OC born before the cutoff	0.0025	0.0003
[Spillover Effect]	(0.0024)	(0.0021)
YC born before the cutoff	0.0119***	0.0119***
[Own Relative Age Effect]	(0.0022)	(0.0023)
Mean(Y)	0.030	0.036
N	106,790	106,790

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. Panel A restricts to cousin pairs in which the older cousin's mother has no college education, while Panel B restricts to cousin pairs in which the older cousin's mother has at least some college education. 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 C16: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's ADHD Drug Treatment in Years 1–3 Following Diagnosis

	(1) Drug in YR1	(2) Drug in YR2	(3) Drug in YR3	(4) No Drug in 3YRs
Panel A: No Covariates				
OC born before the cutoff	0.0023**	0.0015*	0.0014*	0.0011**
[Spillover Effect]	(0.0010)	(0.0008)	(0.0007)	(0.0004)
YC born before the cutoff	0.0098***	0.0089***	0.0067***	0.0016***
[Own Relative Age Effect]	(0.0009)	(0.0008)	(0.0008)	(0.0005)
Mean(Y)	0.030	0.023	0.018	0.006
N	432,903	432,903	432,903	432,903
Panel B: Full Covariates				
OC born before the cutoff	0.0021**	0.0013	0.0012	0.0010**
[Spillover Effect]	(0.0010)	(0.0008)	(0.0007)	(0.0004)
YC born before the cutoff	0.0080***	0.0075***	0.0057***	0.0014***
[Own Relative Age Effect]	(0.0010)	(0.0008)	(0.0008)	(0.0005)
Mean(Y)	0.030	0.023	0.018	0.006
N	432,903	432,903	432,903	432,903

Notes: Each column in each panel reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. The dependent variables in columns (1), (2), and (3) are indicators equal to one if the younger cousin has an outpatient claim with an ADHD diagnosis and obtains an ADHD drug between 0-11, 12-23, and 24-35 months after diagnosis, respectively. The dependent variable in column (4) is an indicator equal to one if the younger cousin has an outpatient claim with an ADHD diagnosis and does not obtain an ADHD drug between 0-35 months after diagnosis. 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 C17: Effect of Older Cousin Being Born Before Cutoff on Younger Cousin's Number of Unique ADHD Drug Treatments and Mental Health Outpatient Visits in Years 1-3 Following Diagnosis

	(1) Unique Drugs YR1	(2) Outpatient YR1	(3) Outpatient YR2	(4) Outpatient YR3
Panel A: No Covariates				
OC born before the cutoff	0.0027**	0.0129***	0.0091***	0.0069**
[Spillover Effect]	(0.0013)	(0.0048)	(0.0027)	(0.0027)
YC born before the cutoff	0.0116***	0.0374***	0.0199***	0.0153***
[Own Relative Age Effect]	(0.0012)	(0.0051)	(0.0030)	(0.0029)
Mean(Y)	0.036	0.132	0.057	0.044
N	432,903	432,903	432,903	432,903
Panel B: Full Covariates				
OC born before the cutoff	0.0023*	0.0120**	0.0087***	0.0064**
[Spillover Effect]	(0.0013)	(0.0048)	(0.0028)	(0.0026)
YC born before the cutoff	0.0094***	0.0311***	0.0174***	0.0136***
[Own Relative Age Effect]	(0.0012)	(0.0052)	(0.0030)	(0.0028)
Mean(Y)	0.036	0.132	0.057	0.044
N	432,903	432,903	432,903	432,903

Notes: Each column reports results from a separate regression estimating model (2). See notes under Table 2 for more information about the sample, specifications, and control variables. The dependent variable in column (1) is the number of unique ATC codes for ADHD drugs obtained in the first year after diagnosis. The dependent variables in columns (2), (3), and (4) are the number of outpatient visits with ICD codes beginning in "F" taking place 0-11, 12-23, and 24-35 months after diagnosis, respectively. For individuals without a diagnosis, these variables are equal to 0. Robust standard errors are clustered on the older cousin's day of birth.

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