Journal of Health Economics 29 (2010) 657–673

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children☆

William N. Evans[a,*], Melinda S. Morrill[b], Stephen T. Parente[c]

[a] Department of Economics and Econometrics, University of Notre Dame, 437 Flanner Hall, Notre Dame, IN 46556, United States
[b] Department of Economics, North Carolina State University, 4144 Nelson Hall, Box 8110, Raleigh, NC 27695-8110, United States
[c] Department of Finance, Carlson School of Management, University of Minnesota, 321 19th Ave. Room 3-122, Minneapolis, MN 55455, United States

ARTICLE INFO

Article history:
Received 8 April 2010
Received in revised form 15 July 2010
Accepted 15 July 2010
Available online 4 August 2010

JEL classification:
I12
I18

Keywords:
ADHD
Stimulants
Regression discontinuity design

ABSTRACT

We exploit the discontinuity in age when children start kindergarten generated by state eligibility laws to examine whether relative age is a significant determinant of ADHD diagnosis and treatment. Using a regression discontinuity model and exact dates of birth, we find that children born just after the cutoff, who are relatively old-for-grade, have a significantly lower incidence of ADHD diagnosis and treatment compared with similar children born just before the cutoff date, who are relatively young-for-grade. Since ADHD is an underlying neurological problem where incidence rates should not change dramatically from one birth date to the next, these results suggest that age relative to peers in class, and the resulting differences in behavior, directly affects a child's probability of being diagnosed with and treated for ADHD.

1. Introduction

Nearly all critics of the U.S. healthcare system note that the U.S. spends far more on health care than any other developed country yet performs poorly in international comparisons on aggregate outcomes such as life expectancy and infant mortality.[1] Some interpret these statistics as an indication that the U.S. health care system is on the "flat-of-the-curve" (Fuchs, 2004) in the health production function meaning the marginal health care dollar is of little or questionable medical value. The notion that a large fraction of health care spending produces little return is bolstered by data from the Dartmouth Atlas which shows that per capita Medicare reimbursements across hospital referral regions vary by a factor of three (Wennberg et al., 2008), yet there is little evidence that these differences in spending lead to better quality of care (Baicker and Chandra, 2004) or better mortality outcomes (Fisher et al., 2003). This same research program suggests that the U.S. could reduce Medicare spending by 30% without any drop in medical outcomes. Similarly, the Institute of Medicine (2007) estimates that nationwide less than half of all treatments delivered are supported by evidence.

☆ Portions of the research presented here were conducted at the Triangle Census Research Data Center. The content of this work does not necessarily represent the official views of the Agency for Healthcare Research and Quality, the Census Bureau, or the National Center for Health Statistics. The authors would like to thank Patricia Anderson, Kelly Bishop, Jason Fletcher, Ezra Golberstein, Julie Hotchkiss, Lisa Kahn, Colleen Manchester, Thayer Morrill, Orgul Ozturk, Sarah Pearlman and seminar participants at the Southern Economic Association Meetings, the Triangle Census Research Data Center Annual Conference, North Carolina State University, UNC Greensboro, UNC Chapel Hill, the PAA Annual Meeting, and AcademyHealth Health Economics Interest Group for helpful comments and suggestions. We especially want to thank Elizabeth Dhuey for sharing her data on school start laws and for numerous fruitful conversations about these state statutes, and to Negasi Beyene, Ray Kuntz and Bert Grider for assistance with the restricted-use NHIS and MEPS data. All remaining errors are our own.

* Corresponding author. Tel.: +1 574 631 7039; fax: +1 574 631 4783.
E-mail addresses: wevans1@nd.edu (W.N. Evans), melinda_morrill@ncsu.edu (M.S. Morrill), paren010@umn.edu (S.T. Parente).

[1] In 2006, per capita spending on health in the U.S. was $6714, more than twice the median value for OECD countries. Despite this spending, in 2005, the U.S. ranked 25th of 29 countries in average life expectancy and the U.S. had the fourth highest infant mortality rate of 28 reporting countries. All data is from the OECD's frequently requested data series, http://www.irdes.fr/EcoSante/DownLoad/OECDHealthData_FrequentlyRequestedData.xls.

0167-6296/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2010.07.005

The statistics reported above have lead to a greater emphasis on reducing waste and improving the quality of clinical decisions as cornerstones of any health care reform initiative. For example, $1.1 billion was earmarked for cost-effectiveness research as part of the American Recovery and Reinvestment Act, signed into law on February 19, 2009 by President Obama.[2]

The difficulty in implementing practice reform is identifying what is and is not medically appropriate. Utilization review is now commonplace in medicine and a large volume of research uses chart review to identify procedures that are appropriately indicated by medical conditions. Unfortunately, chart reviews are expensive and in many instances review can only indicate whether the treatment was appropriate given the diagnosis, not whether the diagnosis itself was correct in the first place.

In this paper, we implement a statistical procedure to examine the medical appropriateness of one specific diagnosis (attention-deficit/hyperactivity disorder) and its most frequent treatment (stimulants). The procedure is implemented using information typically gathered in claims data files or reported in surveys, which greatly reduces the data needs compared to other forms of utilization review.

Attention-deficit/hyperactivity disorder (ADHD) is a neurological condition characterized by delayed brain development (Shaw et al., 2010). According to the National Institute of Mental Health (2008), children with ADHD are hyperactive and tend to have difficulty staying focused and controlling behavior. The ADHD Booklet explains (p. 2) "it is normal for all children to be inattentive, hyperactive, or impulsive sometimes, but for children with ADHD, these behaviors are more severe and occur more often." Not only is ADHD difficult to diagnosis, but often the diagnosis is made by a pediatrician or family physician without consultation with a mental health specialist (Safer and Malever, 2000). In the United States about 5–10% of children aged 6–18 have been diagnosed with ADHD and some estimates suggest this number increased by 500% between the late 1980s and early 2000s (Zuvekas et al., 2006).

In this paper, we provide evidence that the diagnosis and treatment of ADHD is heavily influenced by the relative age of children in school. Most public schools in the United States have an official "age of start" date that indicates the time by which a child must turn 5 years old in order to enter kindergarten. Age at school start laws create quasi-experimental variation in the age of children where those born just before the kindergarten eligibility date may enter school in a given year, while children born only a few days later must wait an entire year to start school. The children born just before the cutoff date are on average younger than their classroom peers. The relative immaturity of these young-for-grade children may be mistaken as ADHD due to the nature of the diagnostic guidelines that suggest a comparison with a child's peers. According to the medical guidelines described by the NIMH ADHD Booklet, health professionals are asked to consider whether the observed behaviors (p. 6) "happen more often in this child compared with the child's peers?" Given age-of-start laws, a typical kindergarten class may contain a child who just turned five and someone almost six, a difference in age of 20%. Using a regression discontinuity model, we exploit the discrete jump in school enrollment generated by kindergarten eligibility laws to examine whether children's relative age influences their probability of being diagnosed with ADHD and, as a result, to be prescribed stimulants.

ADHD is an underlying neurological problem and incidence rates should not change dramatically from one birth date to the next. If diagnosis rates do shift appreciably based on small changes in birth dates, then the diagnosis is not based entirely on underlying conditions. Evidence consistent with increased diagnosis of ADHD for younger children is provided in Elder and Lubotsky (2009) who used samples from the Early Childhood Longitudinal Study – Kindergarten cohort (ECLS-K) data to document persistent negative consequences for younger children in school.

In this paper, we use data on ADHD diagnosis from the 1997 to 2006 National Health Interview Survey (NHIS), plus data on prescription drug use of stimulants from the 1996 to 2006 Medical Expenditure Panel Survey (MEPS), and a nationwide private health insurance company over the 2003 through 2006 time period. In all three samples, we find evidence children whose fifth birthday fell just after the school eligibility cutoff date, who are therefore more likely to be older-for-grade, have significantly lower chances of being diagnosed with, and treated for, ADHD. The effect sizes are large. Children born just after the cutoff date are 2.1 percentage points less likely to be diagnosed with ADHD and 1.6% less likely to be treated with a stimulant, numbers that are roughly 25% smaller than their sample means. As we outline below, the results imply that being young for your grade more than doubles the chance that a student is diagnosed with or treated for ADHD.

The basic results in this paper are quite similar to those in Elder (this issue), who used the same techniques employed here and data from the ECLS-K to demonstrate that children born just before the state's age-of-start cutoff date are 50% more likely to be diagnosed with ADHD than those born just after. The fact that the basic results in this paper can be replicated in four different data sets should reassure those with concerns this finding is spurious.

2. Background on ADHD

According to the National Institutes of Mental Health (NIMH) ADHD Booklet, the characteristic behaviors associated with ADHD are inattention, hyperactivity and impulsivity. These symptoms typically appear early in life and in many cases last into adulthood. Accurate identification of ADHD is critical since children with ADHD are at an increased risk of academic difficulties such as a greater incidence of learning disabilities (Mayes et al., 2000), a higher chance of repeating a grade and lower test scores (Currie and Stabile, 2006), and a higher dropout rate (Trampush et al., 2009). Outside the classroom, children with ADHD have higher rates of illegal drug use (Biederman et al., 1998), greater motor vehicle accident rates (Woodward et al., 2000; Barkley et al., 1993), and a greater likelihood of having other psychiatric conditions (Pliszka, 1998; Jensen et al., 1997). Data from the National Survey of Children's Health indicate that among youths 4–17 years of age, 7.8% reported an ADHD diagnosis with boys having a 2.5 times greater incidence rate than girls (Visser et al., 2007).

Treatments options for children with ADHD include medication management, behavioral treatment, routine community care, or some combination of these regimens. In a random assignment clinical trial, financed by the National Institutes of Mental Health, the Multimodal Treatment Study for Children with ADHD (MTA Cooperative Group, 1999) found a combination treatment of medication management and behavioral treatment and medication management alone produced superior results to behavioral treatment or routine community care.

Despite the variety of treatment options, we focus on prescription stimulant medication for the following reasons. First, stimulants have been demonstrated to be extremely effective at controlling the symptoms of ADHD, but stimulants do not treat the underlying disorder or provide a cure for ADHD. As we document below, stimulants also have a number of potential negative

[2] PL 111-5, http://www.gpo.gov/fdsys/pkg/PLAW-111publ5/pdf/PLAW-111publ5.pdf.

side effects. Finally, prescription medications are easy to identify in standard claims data bases.

Data from the Medical Expenditures Panel Survey indicates that roughly 3% of children under the age of 18 were prescribed stimulants such as Ritalin in 2002, which is roughly five times the prescription rate in 1987 (Zuvekas et al., 2006). Visser et al. (2007) note that in 2003 roughly 55% of children diagnosed with ADHD were taking stimulants. Using data from a large sample of privately insured children, Castle et al. (2007) estimate that by 2005, 4.4% of children aged 0–19 in their sample were using stimulants to treat ADHD, with usage rates increasing by roughly 12% per year over the 2000 through 2005 period. Zito et al. (2000) note a rapid increase in stimulant use among pre-school children.

Perhaps due to this striking increase in the diagnosis and treatment of ADHD, concern has been raised by the medical community, popular press, and parent support groups that this rise may be due to over-diagnosis. There is no pathognomonic marker for ADHD and the intensity of symptoms may fluctuate over time (Angold et al., 2000), making accurate diagnosis of ADHD difficult. Moreover, diagnosis of ADHD is often made without consulting a mental health specialist. Safer and Malever (2000) found that of Maryland public school students taking methylphenidate (i.e., Ritalin) at school 63% had prescriptions from pediatricians, 17% from family practitioners, and only 11% received a prescription from a psychiatrist. Diagnoses are generally made after a medical professional considers a child's behavior in multiple contexts, as reported by the parent, teacher, and child.

Stimulant use and ADHD diagnosis rates vary across groups of similarly defined youths, possibly suggesting that clinical guidelines for diagnosis are not being applied consistently. For example, researchers have found large variation in stimulant use by children across regions of the United States,[3] by race and ethnicity, and by gender.[4] Comparing stimulant use among children in two southeastern Virginia cities, LeFever et al. (1999) found tremendous heterogeneity in stimulant use both within and between cities and conclude that the (p. 975) "criteria for diagnosis of ADHD vary substantially across U.S. populations, with potential over-diagnosis and overtreatment of ADHD in some groups of children." Similarly, in a study of children in the Great Smokey Mountains, Angold et al. (2000) found that the presence of ADHD symptoms is not well correlated with the treatment of ADHD through prescription medication and thus conclude that (p. 135) "stimulant treatment was being used in ways substantially inconsistent with current diagnostic guidelines."

This heterogeneity in diagnosis and treatment rates across gender and race has been documented in many settings. In a large-scale study specifically designed to assess the disparity in treatment, Safer and Malever (2000) collected data on all children that received medication for the treatment of ADHD during school hours in the State of Maryland in 1998. They found that the boys in elementary school were 3.5 times as likely to be receiving treatment as girls, and that black and Hispanic students were about half as likely to be receiving treatment relative to non-Hispanic white students.

Although these studies effectively demonstrate the heterogeneity in diagnosis and treatment rates across different demographic groups, it is difficult to know from these results whether this heterogeneity is a result of genetic or environmental factors, rather than a reflection of inappropriate diagnosis. Because the etiology of ADHD is not well understood, risk factors for ADHD are often based on population averages, such as a being male or having a lower socioeconomic status. While these population averages are somewhat consistent over time and across geographies, there is no clear medical evidence that higher diagnosis and treatment rates are due to a higher prevalence of the disorder in these populations. Comparing diagnosis rates across populations may confound issues such as access to and quality of care for any disease. This is particularly problematic for ADHD diagnosis (and the diagnosis of other mental disorders in childhood) since there is no objective clinical test.

The potential of inappropriate diagnosis and treatment is most troubling when considering the biological effects of the commonly prescribed stimulants. The side effects of methylphenidate use include insomnia, stomachache, headache, dizziness, and decreased appetite (Ahmann et al., 1993). More importantly, stimulants have been shown to increase heart rates and blood pressure (Nissen, 2006). Less is known about the longer-term effects. Because stimulants act to inhibit the dopamine receptors in the brain, there is some concern and speculation that long-term changes in cell function might result from chronic exposure to stimulant medication, particularly during brain development in childhood and adolescence (Volkow and Insel, 2003). In addition to these important medical side effects of stimulant use, there is also an economic cost associated with diagnosis and treatment. Pelham et al. (2007) use a cost of illness framework to estimate the economic impact of ADHD and they conclude that the cost of ADHD is between $12,005 and $17,458 per child in 2005 dollars.

ADHD is often diagnosed after a teacher observes a child in his/her classroom and refers the parent to have the child evaluated. In a survey of physicians in the Washington, DC metro area, Sax and Kautz (2003) found that in 52% of all cases, teachers and other school personnel are the first to suggest a diagnosis of ADHD. It seems natural that teachers should compare the behavior of children within a class and recent research suggests that ADHD diagnosis rates are in fact correlated with the relative age of students within a class. In the most detailed study to date, Elder and Lubotsky (2009) used data from the Early Childhood Longitudinal Study – Kindergarten cohort (ECLS-K) to examine the impact of being older for a grade on a long list of outcomes. The Elder–Lubotsky paper serves as the template for our work in that they use the variation in student age generated by age of school start models to identify their model. Using an instrumental variables framework, the authors found that children who are an additional year of age older at school entry have superior educational outcomes. For example, these older children tended to have higher test scores and fewer behavioral problems. More importantly for our work, the authors demonstrated that starting school later reduces the chance of being diagnosed with ADHD by 50%.

This work is part of a larger literature in labor economics that explores the beneficial cognitive and labor market effects of being among the oldest children in the classroom. Many studies have exploited the variation in school start eligibility laws across states, over time, and even between countries.[5] For example, using international data, Bedard and Dhuey (2006) demonstrated that being young for your class produces lower test scores through the eighth grade. More recently, Dhuey and Lipscomb (forthcoming) find that

[3] Cox et al. (2003) demonstrated tremendous regional variation in stimulant use in a sample of children with private insurance.

[4] Castle et al. (2007) found that boys ages 0–19 were 2.3 times more likely to receive stimulant medications than girls in a comparable age range for a sample of children in a private prescription claims database. Visser et al. (2007) found gender and race/ethnicity are related to ADHD diagnosis, but not to ADHD medication treatment.

[5] See, for example, Angrist and Krueger (1992), Bedard and Dhuey (2006), Datar (2006), Elder and Lubotsky (2009), Dobkin and Ferreira (2010), Fertig and Kluve (2005), Goodman et al. (2003), Lincove and Painter (2006), McEwan and Shapiro (2008), Puhani and Weber (2007), and references therein.

relative age in the classroom causes a higher risk of being labeled as having a learning disability. Given this large literature on age effects and given the stark change in ADHD diagnoses rates based on age of school start found in the Elder and Lubotsky paper, we also suspect a similar disparity in stimulant use rates. In this paper, we replicate the basic results in Elder and Lubotsky (2009) using restricted-use data from the National Health Interview Survey and state data on age of school start legislation. We then extend these basic models to include data on stimulant use.

While completing the work for this paper, we came across the independent work of Elder (this issue), who used the same techniques employed here and data from the ECLS-K to demonstrate that children born just before the state's age-of-start cutoff date are 50% more likely to be diagnosed with ADHD than those born just after. The robustness nature of the results across samples in this paper and the work of Elder is encouraging and suggests that the results presented below are not spurious but represent true misdiagnosis of ADHD.

3. Empirical specifications

The primary question we consider is whether children that are older for their grade are less frequently diagnosed with and treated for ADHD. A similar set of questions has been addressed in a variety of disciplines about whether delayed entry into school helps or hinders academic promise. The underlying structural equation for both questions is essentially the same. Let the unit of observation be the individual child, indexed by i, and let Y_i be a dummy variable that equals 1 if a student is diagnosed (or treated) for a particular condition such as having developmental problems. The focus of this paper is ADHD and therefore, in our context, Y would equal 1 if a child is diagnosed (or treated) for ADHD. A student is defined as young for their grade (Young_i) if they are below some threshold age such as the median for children in the same state, grade, and year. The primary equation of interest is therefore

$$Y_i = \beta_0 + x_i\beta_1 + \text{Young}_i\beta_2 + h(z_i) + \kappa_i \quad (1)$$

where x is a vector of observed characteristics and κ is a random error. The function $h(\cdot)$ is a smooth function in z, a variable that measures the difference in days between the child's birth date and the state cutoff date when that child was age five. Given a state with a September 1st age at start cutoff, a September 1st birth date would have a value of $z = -1$, a September 2nd birth date would be $z = 0$ and an October 1st birth date would be have a value of $z_i = 29$. Following previous applications, we capture $h(z)$ with polynomial terms in z and interactions of these polynomials with the treatment indicator $I(z_i \geq 0)$.

If children of different ages were randomly assigned to classes, ordinary least squares (OLS) estimates of the parameter of interest (β_2) would be consistent. There is, however, good reason to suspect that single-equation estimates of Eq. (1) are subject to an omitted variables bias. Parents often decide their child is not ready for kindergarten and enroll their child in school later than others from the same birth cohort. This behavior is often referred to as "academic redshirting." If parents delay a child's entrance into kindergarten because they have difficulty sitting still or focusing on school work, which in turn signals a greater likelihood of an ADHD diagnosis in the future, then redshirting signals reverse causation from diagnosis to age relative to peers and OLS estimates of Eq. (1) would then understate the coefficient on β_2.

The available evidence suggests this is a real concern. West et al. (2000) estimate that during the mid 1990s, roughly 9% of students delayed entry into kindergarten. Males were 30% more likely than females to have delayed entry and children with diagnosed development problems were more than twice as likely as those without such diagnoses to have delayed entry. The number of academic redshirts and the role that developmental issues play in the decision suggests that estimating Eq. (1) by OLS will lead to inconsistent and potentially misleading estimates of the effect of relative age on ADHD diagnosis.[6]

We could obtain a consistent estimate for β_2 if we could somehow mimic random assignment and alter the relative ages of children in classes in a way that conveys no direct information about underlying ADHD incidence. In just this fashion, we use the distance between a child's birthday and the age at school entry as an instrument for relative age in class within a regression discontinuity design (RDD) or an instrumental variables (IV) model.

Children born a few days apart should be, on average, similar along all characteristics (e.g., underlying intelligence, parental backgrounds, home environment, etc.) yet because of age of school start laws these children will have vastly different ages when they start school. Consider a state that has a September 1st cutoff date. In this state, children born on August 31st are more likely to begin school as a 5 year old, but those students born just a few days later, on September 2nd, must wait a year to begin school. This age difference in a class is relatively large in early grades. Around the start of the school year, a class containing students with an August 31st and a September 2nd birth date will differ in age by 20% in kindergarten, 14% in second grade and 10% in fifth grade. The sharp break in age at school start generated by the interaction of child birthdates and the assumed similarity of children born just before and just after the school cutoff suggests that any observed difference in ADHD diagnosis and treatment between these two groups can be attributed to the difference in ages of the children in school.

Instrumental variables (IV) estimates of Eq. (1) can be obtained in two steps. The initial step is to examine the first-stage relationship between age relative to the state cutoff date and the relative age in class. This model can be represented by the equation

$$\text{Young}_i = \gamma_0 + x_i\gamma_1 + \gamma_2 I(z_i \geq 0) + h(z_i) + \nu_i, \quad (2)$$

where $h(z)$ and x are defined as above, ν is a random error and the dummy variable $I(z_i \geq 0)$ equal 1 if the student has a birth dates after the age at school start. The impact of the age at start laws on whether the child is young for the class is captured by the parameter γ_2. The key assumption of the RDD model is that in the absence of the treatment (in this case, the student's birth date occurs after the school start cutoff) the outcome of interest is "smoothly" changing in z (the child's age) which is captured by the polynomial $h(z)$. Given $h(z)$, we assume that people on either side of z_i in the absence of age of start laws are functionally identical, controlling for observable characteristics x.

The second step in the process is to examine the reduced-form relationship between a child's age relative to the school start dates and their diagnosis and/or treatment of ADHD. This relationship can be captured by the following equation

$$Y_i = \alpha_0 + x_i\alpha_1 + \alpha_2 I(z_i \geq 0) + h(z_i) + \zeta_i \quad (3)$$

where ζ_i is a random error and all remaining variables are defined as above. Given the assumptions above, the coefficient α_2 measures the impact of being born just after the cutoff on the propensity of students to experience the outcome Y_i. Because this is an exactly identified model with one endogenous variable, the IV estimate of β_2 in Eq. (2) is obtained by simply dividing α_2, the impact being born after the age of start on ADHD diagnosis, by the fraction of

[6] Despite these concerns, many such models have been estimated in the past (Byrd et al., 1997; Stipek and Byler, 2001; Lincove and Painter, 2006).

people impacted by the age of start (γ_2 from Eq. (2)), or

$$\hat{\beta}_2 = \hat{\alpha}_2/\hat{\gamma}_2. \quad (4)$$

Arithmetically, this is also equivalent to estimating Eq. (2) by two-stage least-squares (2SLS) and using $I(z_i \geq 0)$ as an instrument for Young$_i$.

The difficulty with Eq. (4) is that our data are not well suited for estimating the first-stage model outlined in (2). As we describe below, our large sample of private claims data, which measures drug use to treat ADHD, does not contain data on a child's current grade. The two nationally representative samples, the National Health Interview Survey (NHIS) and the Medical Expenditure Panel Survey (MEPS), ask respondents for the highest grade completed, which requires that we impute current grade by adding one to the recorded value for children currently enrolled in school. This is problematic for two reasons. First, we will overstate current grade for those who have completed but must repeat a grade. Second, it appears that some parents are reporting the child's current grade rather than the highest grade completed, meaning that by imputing the grade, we will have too many respondents that are young for their class.

To verify this point, we extracted a sample of children aged 7–16 from the 2000 to 2002 October Current Population Survey (CPS) data sets. These data contain a school enrollment supplement that identifies the current grade enrolled for all respondents. In Appendix Fig. 1, we report the distribution of grades relative to age for this sample. Almost 70% of students are in a grade that is 5 years lower than their age (most 8 year olds enrolled in school in October are in the third grade) with the next largest group enrolled in a grade that equals age minus six, and a few students are young for their grade, enrolled in a grade that is 4 years lower than age.

We compared these numbers to those who responded to the NHIS in the fourth quarter of the year. For this sample, we take data from the 2000 to 2002 NHIS, and use reported month and year of birth to impute the respondent's age as of October 1st to make this sample as comparable as possible to the October CPS. We add 1 year to the highest grade completed in order to estimate the current grade enrolled. Graphing the implied distribution of grades for age from this sample in Appendix Fig. 1, we see that the NHIS overstates by a factor of three the number of students that are young for their class (in grade = age − 4) and understates by 40% the fraction who are older for their class (in grade = age − 6).

In practice, the systematic measurement error in the imputed current grade from the NHIS will tend to understate the first-stage coefficient γ_2, which will overstate the implied IV estimate in Eq. (4). For this reason, we will rely more on the reduced-form models in Eq. (3) to signal the causal relationship between being young for class and ADHD diagnosis and treatment than on the IV estimates.

There is both between-state variation in the age at school start and within-state variation in these laws over time.[7] A summary of the cross-sectional and time series variation in these laws is shown in Table 1. Seven states (CO, MA, NH, NJ, NY, PA, and VT) had no statewide age at school entry law in 2005, but rather allowed local education authorities (LEA) to determine age at school entry standards. Twenty-five states (including the District of Columbia) have had the same age at school start date since 1984, while the rest have had changes at some point in the period. In the 2005/2006 school year, the age at school start cutoff dates vary anywhere from July 1st in IN until January 1st in CT.

4. Data

The data requirements for the RDD model outlined above are substantial. Naturally we need a data set that identifies whether a child has been diagnosed with ADHD and/or whether that child uses a prescription stimulant medication to treat ADHD. In addition, we must identify a child's exact date of birth and state of residence so that we can calculate his/her age relative to the kindergarten eligibility cutoff date. These last set of descriptors are identifying variables that are not typically available on public use versions of data sets. Consequently, we estimate the empirical models on three separate restricted-access data sources: the National Health Interview Survey (NHIS), the Medical Expenditures Panel Survey (MEPS), and a private insurance prescription drug claims data set. Even though our data cover different time periods and populations, we find similar results in each data set, confirming the robustness of our findings.

The NHIS is an annual survey of roughly 60,000 households that collects data on the extent of illness, disease, and disability in the civilian, non-institutionalized population of the United States. The NHIS includes detailed demographic and socioeconomic information, as well as the self-reported medical conditions of respondents. Information on ADHD diagnosis has been included in the Sample Child Supplement within the NHIS since 1997. Our empirical strategy relies on the ability to identify the exact cutoff date that each child faced when they first entered kindergarten, plus their birth date. We therefore use the more detailed geographic data and the exact date of birth that is available only in the restricted-use version of the NHIS.[8] The dependent variable for the NHIS analysis is the child's parent's report of whether the child has ever been diagnosed with ADHD by a doctor or health professional. ADHD incidence rates from the NHIS are comparable to results from other national surveys from similar periods.

Our second data source, the Medical Expenditure Survey (MEPS), is a series of surveys administered since 1996 by the Agency for Healthcare Research and Quality and the National Center for Health Statistics. The MEPS sample is drawn from the NHIS sample, although there are restrictions on merging these two datasets. There are three components of the survey completed by households, medical providers, and insurance companies. Individuals are asked questions over a series of five rounds detailing 2 years of medical expenditures and services utilization. Each year of the MEPS contains respondents from two overlapping panels. The MEPS full-year consolidated data file (CDF) contains sociodemographic information for respondents including age, sex, race, and basic economic characteristics, plus their date of birth. We have access to the restricted version of the MEPS, which allows us to identify the exact eligibility data as described above.[9] While the MEPS is a smaller sample than our private claims data, as with the NHIS, it has the advantage that it contains children with any health insurance type, including those that are uninsured. The dependent variable for this part of the analysis is whether a child receives a prescription for a stimulant to treat

[7] For a discussion of the individual state statutes and a detailed breakdown of the age of school entry laws in the U.S. from the early 1980s through the present time, see Morrill (2008).

[8] These data are available for use through the National Center for Health Statistics Research Data Center: http://www.cdc.gov/nchs/r&d/rdc.htm. We access the data at the Triangle Census Research Data Center through a data sharing agreement made between the Census Bureau and the National Center for Health Statistics.

[9] The restricted access MEPS is available at regional Research Data Centers through a data sharing agreement made between the Census Bureau and the Agency for Healthcare Research and Quality. We access the data at the Triangle Census Research Data Center.

Table 1
Kindergarten eligibility cutoff dates.

State	Cutoff 2005	Law changes since 1984	State	Cutoff 2005	Law changes since 1984
AL	1-Sep	1984–1989: 10/1	MD	30-Sep	1984–2002: 12/31
		1990+: 9/1			2003: 11/30
AK	1-Sep	1984–1987: 11/2			2004: 10/31
		1988–2003: 8/15			2005: 9/30
AZ	31-Aug*				2006+: 9/1
AR	15-Sep	1984–1997: 10/1	MA	LEA	
		1998: 9/1	MI	1-Dec	
		1999+: 9/15	MN	1-Sep	
CA	2-Dec	1984–1986: 12/1	MS	1-Sep	
		1987+: 12/2	MO	31-Jul*	1984–1986: 8/31*
CO	LEA				1987: 7/31*
CT	1-Jan				1988–1996: 6/30*
DE	31-Aug	1984–1992: 12/31			1997+: 7/31*
		1993: 11/30	MT	10-Sep	
		1994: 10/31	NE	15-Oct	
		1995: 9/30	NV	30-Sep	
		1996+: 8/31	NH	LEA	
DC	31-Dec		NJ	LEA	
FL	1-Sep		NM	31-Aug*	
GA	1-Sep	Established 1985	NY	LEA	
HI	31-Dec		NC	16-Oct	
ID	1-Sep	1984–1989: 10/16	ND	31-Aug*	
		1990: 9/16	OH	30-Sep	
		1991–1992: 8/16	OK	1-Sep	
		1993+: 9/1	OR	1-Sep	1984–1985: 11/15
IL	1-Sep	1984–1985: 12/1	PA	LEA	
		1986: 11/1	RI	1-Sep	1984–2003: 12/31
		1987: 10/1	SC	1-Sep	1984–1992: 11/1
		1988+: 9/1	SD	1-Sep	
IN	1-Jul	1984–1988: LEA	TN	30-Sep	1984: 10/31
		1989: 9/1	TX	1-Sep	1984–1994: ssy
		1990: 8/1			1995+: 9/1
		1991: 7/1	UT	1-Sep*	1984–1987: ssy
		1992–2000: 6/1			1988+: 9/1*
		2001–2005: 7/1	VT	LEA	1984–1990: 1/1
IA	15-Sep				1991+: LEA
KS	31-Aug	1984–1994: 9/1	VA	30-Sep	
		1995+: 8/31	WA	31-Aug	
KY	1-Oct		WV	31-Aug*	
LA	30-Sep	1984–1995: 12/31	WI	1-Sep	
		1996+: 9/30	WY	15-Sep	
ME	15-Oct				

Notes: Data acquired from individual state statutes. LEA denotes that the state allowed the local education authority to determine the applicable cutoff, therefore there is no statewide date. Starred dates indicate that the statute specifies that the child must be born *before* a certain date, so we have adjusted the date in this table to reflect the date that the child must be born on or before to be consistent across states. ssy: start of school year

ADHD in the survey year. We pool observations across survey years, so a subset of children are represented in the dataset twice. We rely on the ICD-9 codes that identify whether the child received any medication for the treatment of ADHD (ICD-9 code 314).[10]

We have also obtained a proprietary claims data base constituting private insurance contracts for nearly 1 million covered lives and representing at least 40 of the 50 U.S. states. The data set contains claims and health insurance enrollment data for the 2003 through 2006 years of service. The data provide specific information on an insured's date of birth, age, gender, zip code of residence, insurance contract type (e.g., single, two person, family) and premium paid by the insured. Claims data elements of interest include date of service, ICD9 diagnosis and CPT4 procedure code (if medical care) and NDC drug code (if pharmacy). In addition, the pharmacy data provides information on days of supply and refill rates. Both medical and pharmacy data describe the amount paid by the insurer as well as the insured. All of the insured ID information has been encrypted and stripped of any identifying information.[11]

When using the private claims data, the dependent variable is whether in a given year the child had a claim for a prescription drug that is typically used to treat ADHD. Although Ritalin is the most common drug prescribed to treat ADHD, there are many drugs on the market and in recent years, several new drugs have been devel-

[10] Note that we only include medication that was not imputed and that was recorded as being for the primary diagnosis of ADHD. Relaxing these two restrictions increases the mean rate of treatment but does not affect the qualitative conclusions from the regression results.

[11] Because the encrypted Social Security number was missing for a number of dependent children, we could not use that variable to uniquely identify respondents in this sample. Instead, we used the employee's encrypted Social Security number and the dependent's date of birth, which necessitated that we delete twins and higher parity births from the sample. Our results are not sensitive to this restriction.

Table 2
Sample characteristics.

Variable	Full sample	Eligible states	Regression sample		
			±120 Days	Born before [−120, −1]	Born after [0, 120]
National Health Interview Survey (NHIS) [1997–2006]					
Observations (person/year)	69,350	53,212	35,343	17,728	17,615
% Male	51.0%	50.8%	50.7%	50.3%	51.0%
Average age as of June 1	11.9	11.8	11.8	11.7	11.8
% White (Non-Hispanic)	64.6%	64.5%	64.2%	64.1%	64.4%
% Black (Non-Hispanic)	15.4%	16.0%	15.9%	16.0%	15.9%
% Hispanic	15.7%	15.0%	15.2%	15.1%	15.2%
% Other race/ethnicity	4.4%	4.5%	4.7%	4.8%	4.5%
% ADD/ADHD diagnosis	8.4%	8.7%	8.6%	9.7%	7.6%
Medical Expenditure Panel Survey (MEPS) [1996–2006]					
Observations (person/year)	59,814	47,423	31,641	15,952	15,689
% Male	51.1%	51.1%	51.2%	50.9%	51.5%
Average age as of June 1	12.0	11.9	11.9	11.9	11.9
% White (Non-Hispanic)	62.5%	62.6%	62.4%	62.0%	62.7%
% Black (Non-Hispanic)	15.7%	15.9%	16.0%	15.9%	16.1%
% Hispanic	16.3%	15.7%	15.5%	15.8%	15.3%
% Other race/ethnicity	5.5%	5.9%	6.1%	6.2%	6.0%
% Any stimulant use	4.2%	4.3%	4.3%	4.5%	4.0%
Private claims data [2003–2006]					
Observations (person/year)	121,352	72,885	48,206	24,380	23,826
% Male	50.3%	50.2%	50.2%	50.3%	50.1%
Average age as of June 1	12.3	12.4	12.4	12.4	12.4
% any stimulant use	5.2%	5.6%	5.8%	6.5%	5.2%

Notes: Data are from the restricted-access versions of the 1997–2006 National Health Interview Survey (NHIS), the 1996–2006 Medical Expenditure Panel Survey (MEPS), and a private insurance claims dataset. The NHIS and MEPS statistics utilize the survey sample weights. The full sample includes children ages 7–17 on June 1st of the survey year. The eligible sample includes children who live in states with clearly defined kindergarten eligibility cutoff dates in the state they reside in the year they turned 5 years old. The regression sample restricts this group to children whose birthdays are within 120 days of school start.

oped to treat this condition. We identify stimulants through the National Drug Codes (NDC) which are 10-digit, 3-segment numbers that identify the manufacturer, item and size/type, respectively. The list of stimulants includes popular drugs such as Ritalin, Metadate, Methylin, Daytrana, and Concerta (methylphenidate), Adderall (amphetamine and dextroamphetamine), and Dexedrine (dextroamphetamine).

We do not pool these three datasets together, but rather present estimates from each separately. The NHIS includes a measure of diagnosis only. The private claims data only measure prescriptions, not diagnosis. The MEPS data also measure prescriptions, but for a nationally representative sample that is not directly comparable to the private claims sample. Because all three of our data sources have significant restrictions on accessing the data and reporting statistics, it is not possible to combine them. In order to ensure that the children in our samples are currently enrolled in school, in all three samples we restrict our attention to children ages 7–17. Most states require that children ages 7–17 be enrolled in school full-time. We also limit the sample to those observations where there was a state-wide age at school start law in force when the child was 5 years of age. We include in the sample only children born within 120 days of the school eligibility cutoff date in their state and year. The final estimation sample used from the NHIS includes 35,343 children. The final sample size from the MEPS is 31,641 observations representing 18,559 children.

Given the geographic distribution of the insurance company and eliminating states with no age at school start law and states with less than 200 person/year observations, in the private claims data we are left with 48,206 observations from 32 states representing data for 22,371 children aged 7–17. Although these data are for individuals with private health insurance and therefore are not nationally representative, having a sample this large enables us to obtain precise estimates and to explore potential heterogeneity in the effects across gender and age. One limitation of the claims data relative to our other two datasets is the lack of demographic information outside of gender and age. However, as we indicate below where we test the sensitivity of our results to the inclusion of a richer set of covariates in the MEPS and NHIS samples, because people born just before and after the age of school start dates are similar on observed dimensions, the addition of demographic controls does not materially alter the statistical results.

Note that ideally we would like to have information on what state the child was residing in during the fall of the year they turned five. We do not have this information in any of the three data sources. In all three we do observe the current state of residence. The NHIS also includes the child's state of birth.[12] In the empirical section we present results that confirm estimates are not sensitive to using state of birth rather than state of residence or restricting to children who reside in the same state in which they were born. Not having state of residence at age five is not a significant limitation for two reasons. First, there is little cross-state movement among school-aged children. In a sample of children age 6–18 from the 2000 Census One-Percent Public Use Micro Samples (PUMS), only 7.7% moved across state lines in the past 5 years.[13] Interstate moves will only contaminate the analysis if they occur differentially for children born just before or after the age of start cutoff or if they occur differentially for children with ADHD. We have some information on the former concern in that data from the 1980 Census One-Percent PUMS indicates that there is little variation in within state moves based on a child's quarter of birth. In that sample, we estimate that among children 6–18 years of age, the fraction that moved in the past 5 years for those born in quarters 1 through 4 are 4.5, 4.7, 4.7 and 4.5%, respectively.[14] The small fraction of children that move after they start school and the lack of large variation

[12] In the NHIS, approximately 10 percent of the sample is missing state of birth. Of those that have both state of residence and state of birth, approximately 10 percent report being born in a different state than they currently reside.
[13] Author's calculations from the Census PUMS files, see Ruggles et al. (2010).
[14] Author's calculations from the Census PUMS files, see Ruggles et al. (2010).

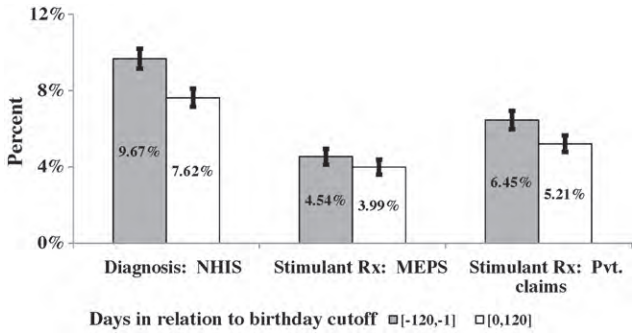

Fig. 1. Means and 95% confidence intervals for children born before and after cutoff dates.
Notes: The means for children born before (dark color) versus after (light color) the kindergarten eligibility cutoff date in their state of residence is shown for children born within 120 days of the cutoff date. Data are from the restricted-access versions of the 1997–2006 National Health Interview Survey (NHIS), the 1996–2006 Medical Expenditure Panel Survey (MEPS), and a private insurance claims dataset. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff. The NHIS and MEPS results are weighted means.

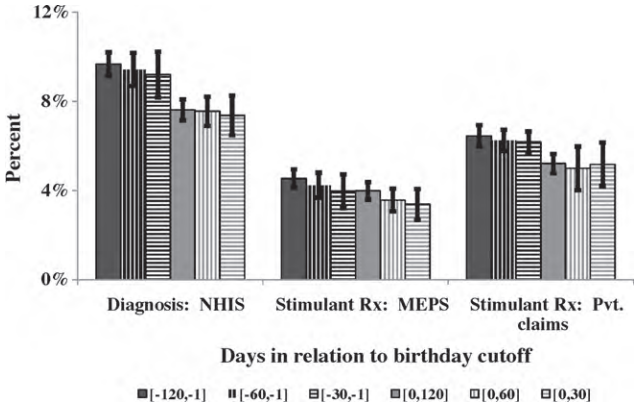

Fig. 2. Means and 95% confidence intervals for children born before and after cutoff dates.
Notes: The means for children born before (dark color) versus after (light color) the kindergarten eligibility cutoff date in their state of residence is shown for children born within 120 (solid), 60 (vertical stripes), and 30 (horizontal stripes) days of the cutoff date. Data are from the restricted-access versions of the 1997–2006 National Health Interview Survey (NHIS), the 1996–2006 Medical Expenditure Panel Survey (MEPS), and a private insurance claims dataset. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff. The NHIS and MEPS results are weighted means.

across birth quarters suggest that using state of residence should not contaminate our results.

5. Results

Table 2 reports sample means and descriptive statistics for each of the three different data sets. In each case, we begin with a sample of children aged 7–17 on June 1st of the survey year. We call this our full sample. Although incidence rates vary by gender, we begin by initially pooling results for males and females. Next, to create the regression sample, we first restrict each sample to children who live in states with a clearly defined kindergarten eligibility cutoff date.[15] Table 2 demonstrates the effect of restricting the sample in this way. While the percent male and average ages are identical in the full and eligible state sample, there is a slightly higher incidence of ADHD diagnosis and treatment in the states used for analysis. As was discussed in Section II, this is consistent with the geographic variation in ADHD treatment and diagnosis rates widely documented in the literature. Next we further restrict the sample to children whose birth date is within 120 days of the cutoff date. While this effectively removes one-third of the sample, we find that the regression sample is very similar to the eligible states sample in each data set. Note that because the private claims data are from a later time period and are, by definition, for a sample of children with private health insurance, we find higher rates of stimulant use than in the MEPS.

The last two columns of Table 2 demonstrate the basic relationship hypothesized above when comparing the fraction of children with an ADHD diagnosis for those born before the cutoff date (students are on average young-for-grade) and children born just after the cutoff date. Notice that in all three data sets the samples of children born just before the cutoff date have nearly identical demographic characteristics when compared with children born just after the cutoff date. However, we find large differences in ADHD diagnosis and treatment rates. In the NHIS, children born before the cutoff experience a 9.7% diagnosis rate compared with only 7.6% for those born after. Stimulant usage in the MEPS indicates a 0.5 percentage point difference between children born before and children born after the cutoff date. Similarly, in the private claims data the percentage of children with any stimulant use drops from 6.5 to 5.2% across the kindergarten eligibility cutoff date.

Fig. 1 presents the graphical equivalent to the means presented in Table 2 and described above. The error bars in the graph repre-

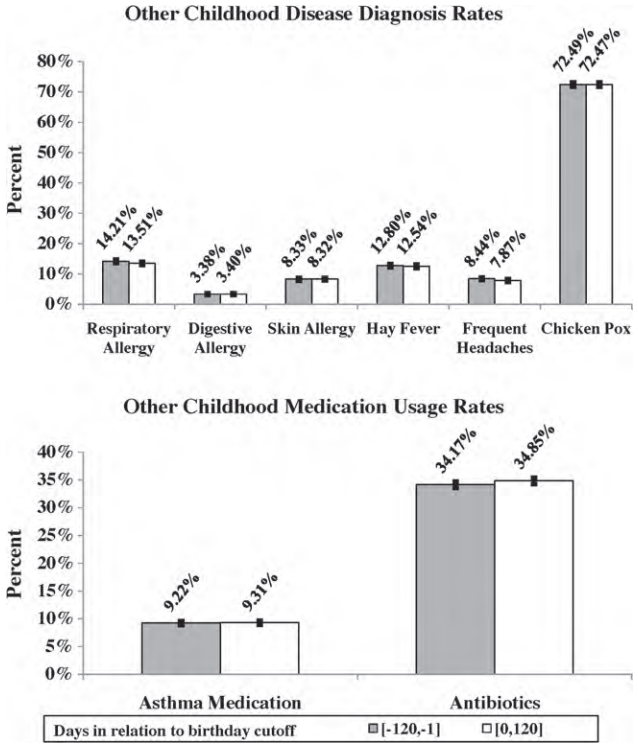

Fig. 3. Falsification tests, means and 95% confidence intervals for children born before and after cutoff dates.
Notes: Childhood disease data are from the restricted-access National Health Interview Survey (NHIS) 1997–2006. Childhood medication use data are from a private claims data set. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff. For the NHIS data, all means are weighted.

[15] Data confidentiality restrictions prohibit the delineation of which states are included in these tables. We have a large enough sample from many states and years to assure a reasonably representative population.

Table 3
First-stage estimates of the effect of being born after the cutoff date on being younger-than-median age in grade × state × year.

Covariates	Models			
	(1)	(2)	(3)	(4)
All Children, 1997–2006, $N = 34{,}173$				
Born after cutoff	−0.3840 (.0224)	−0.3733 (.0221)	−0.3744 (.0220)	−0.3544 (.0186)
Age fixed effects, gender, race/ethnicity		×	×	×
State and birth cohort fixed effects			×	×
1st order polynomial				×
Privately insured children, 2003–2006, $N = 7987$				
Born after cutoff	−0.3710 (.0275)	−0.3604 (.0264)	−0.3615 (.0262)	−0.3565 (.0355)
Age fixed effects, gender, race/ethnicity		×	×	×
State and birth cohort fixed effects			×	×
1st order polynomial				×

Notes: Data are from the 1997 to 2006 National Health Interview Survey, and the sample is restricted to children born ages 7–17 on June 1st of the survey year who were born within 120 days of the kindergarten eligibility cutoff. Coefficients are from linear probability regressions with standard errors in parentheses. Population weights are used and the standard errors are clustered by state.

sent 95% confidence intervals around the sample means. We see that the difference in ADHD diagnosis and treatment rates is large for all samples in all three data sets.[16] In Fig. 2 we present means for progressively smaller samples of children, those born within 120, 60, and 30 days of the kindergarten eligibility cutoff date, respectively. Note that the NHIS is measuring diagnosis, while the MEPS and private claims data include only children receiving prescription stimulants to treat ADHD.

Fig. 3 presents a similar design using six different common childhood ailments found in the NHIS and two other classifications of drugs in the private claims data. The pattern shown in Fig. 1 is unique to ADHD; there is no statistically significant difference in means across kindergarten eligibility cutoff dates for any of these other childhood diseases and other common children's prescription medications.

As discussed in Section III, we caution that IV estimates outlined in Eqs. (1)–(4) may be systematically biased up because of the persistent measurement error in the NHIS education variable. However, in Table 3 we provide an indication of the basic first-stage relationship in the NHIS data. In the top half of the table, we report results for all children aged 7–17. We first define the child's grade as of January 1st of the interview year. To do this, we add one to the last grade completed for those interviewed in the first or second quarter.[17] We then drop observations where the grade level is greater than 12 or where the grade is more than 3 years from the age-appropriate grade level.[18,19] Given these restrictions, the sample size for Fig. 4 is 34,173 children. As in Eqs. (1) and (2) in Section III, we define *Young* as an indicator for whether the child is below the median age in her grade by state by year cell. In Table 3, we report the coefficient on the indicator $I(z_i \geq 0)$ for various specifications. Moving from column 1 to 4, we add progressively more covariates. In column 1, we only include the dummy variable $I(z_i \geq 0)$. Next, in column 2, we add age fixed effects, a

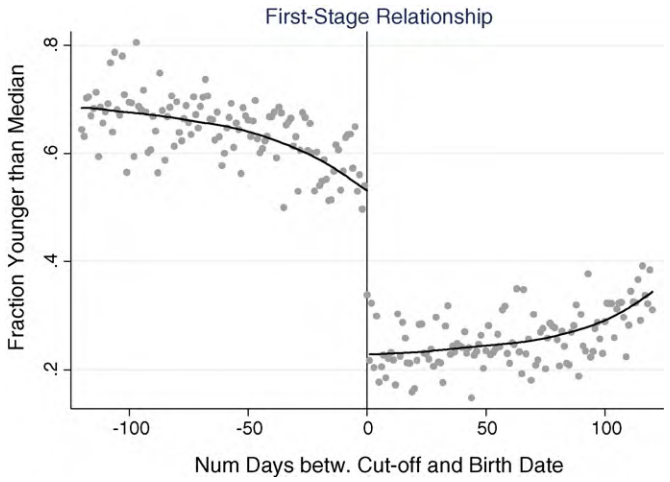

Fig. 4. Fraction of children younger than median for state × grade × year cell by days from kindergarten eligibility cutoff date.
Notes: Data are from the restricted-access versions of the 1997–2006 National Health Interview Survey (NHIS). The horizontal axis indicates bins for children born each number of days from the kindergarten eligibility cutoff date. The dots are the fraction of children in that bin that are younger than the median age for their grade × state × year cell. The lines are from locally weighted regression interpolation. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff.

dummy indicator for male, and controls for race and ethnicity.[20] Column 3 includes a complete set of state and birth cohort effects. Finally, in our preferred specification reported in column 4, we add a first-order polynomial in z with separate trends included for days before ($I(z_i < 0)$) and for days after ($I(z_i \geq 0)$). In all models, we allow for arbitrary correlation in the errors within a state.

Moving from column 1 to 4 in the top half of Table 3, the coefficient on $I(z_i \geq 0)$ falls in absolute value from −0.38 to −0.35. These results suggest that being born just after the cutoff decreases the probability a student is below median in age by 35 percentage points. In all cases, the standard errors are very small and we can easily reject the null that the coefficients are zero at conventional levels. As we report in Table 2, the mean demographic characteristics do not differ across the cutoff date, so it not surprising that adding covariates to the model does not significantly affect the coefficient of interest.

[16] The differences are statistically significant, results available upon request.
[17] In quarter 2 we only add 1 year if the interview month is May or earlier (when available) or assignment week 9 or earlier (when available). The results are not sensitive to these adjustments. The difficultly in determining when the school year would have ended, and thus when the "last grade completed" is equal to the "grade level on January 1st," illustrates the larger problem that grade level is not well measured in the NHIS.
[18] Recall the sample consists of children ages 7–17, where age is defined as the child's age on June 1st of the survey year. The age range allowed in each grade is: Grade 1 (Age 7–9), Grade 2 (Age 7–10), Grade 3 (Age 7–11), Grade 4 (Age 7–12), Grade 5 (Age 7-13), Grade 6 (Age 8-14), Grade 7 (Age 9-15), Grade 8 (Age 10–16), Grade 9 (Age 11–17), Grade 10 (Age 12–17), Grade 11 (Age 13–17), and Grade 12 (Age 14–17).
[19] Note that the measure of age relative to median uses age measured in days.
[20] We define four categories for race/ethnicity: non-Hispanic white, non-Hispanic black, Hispanic, and other. Both the NHIS and MEPS include variables with recoded race and Hispanic origin.

In Fig. 4, we provide a graphical treatment of the results in Table 3. In this graph, the horizontal axis reports days in relation to the cutoff and the vertical axis is the fraction of students who are young for their class. Each dot is a cell mean and the solid line is from a locally weighted regression smoother. The graph clearly shows that children born after the kindergarten eligibility cutoff date in their state by year are considerably less likely to be young for their grade with the difference being approximately 35 percentage points.

If compliance with the kindergarten eligibility cutoff dates were perfect and if grade level was perfectly measured, the coefficient on *born after* should be −1. The coefficient will fall below 1 in absolute value if parents choose not to enroll an eligible child or apply for a waiver to allow an ineligible child to enter early. Likewise, compliance is reduced if children are either held back or advanced a grade. This pattern is found in other work. These choices would result in the instrument, *born after*, having less predictive power for relative age, *Young*. Using data from the Early Childhood Longitudinal Study (ECLS) and the National Education Longitudinal Study (NELS), Bedard and Dhuey (2006) found that in the United States relative age (birth month relative to the school cutoff date) predicts the observed age. For the sample of 4th graders from the ECLS, they found a coefficient of 0.774, while the 8th grade sample from the NELS had a coefficient of only 0.438. These results suggest that compliance with the cutoff date declines as children age, potentially due to grade retention or promotion policies.

Imperfect compliance or strategic behavior on the part of parents should not, however, bias our results. As the compliance weakens the first-stage, the impact of being born after the cutoff on ADHD diagnosis and treatment should also fall, reducing the size of the reduced-form coefficients as well. Because we are only estimating the effects for children whose school entry age is affected by the cutoff date, the smaller first-stage simply means we are defining the treatment effect over a subset of the total population.

In the lower half of Table 3, we report estimates for the first-stage relationship for a sample of children in the 2003–2006 period that had private health insurance.[21] This is a sample that roughly corresponds to the group associated with our private claims sample. The samples do not overlap completely since the private claims data do not include all states in the NHIS. Nonetheless, in the column 4 model that includes a detailed set of covariates, the estimates for the full NHIS sample and the restricted private insurance sample have the same coefficient on $I(z_i \geq 0)$ out to two decimal places. Again, in this more restrictive sample, the standard error on the first-stage estimate is very small.

Next, in Fig. 5 we present a graphical display of the reduced-form model in the NHIS, namely, the impact of being born after the cutoff on being diagnosed with ADHD. In this figure, we see around a 2 percentage point difference in incidence rates between those children that were born just before the cutoff date when compared to those born just after, which is about 25% of the sample mean.

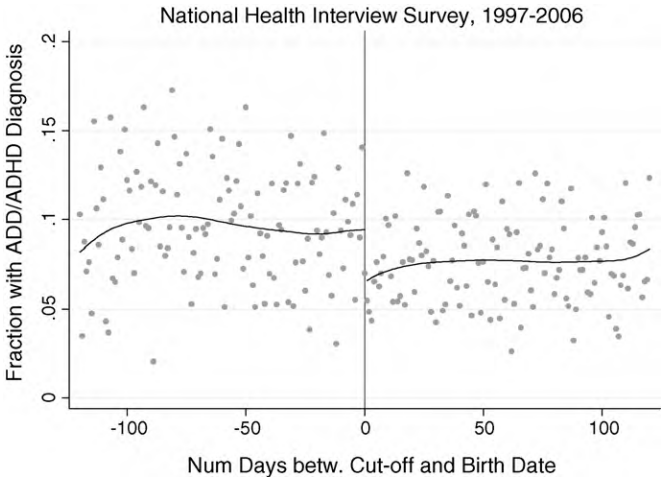

Fig. 5. ADHD diagnosis by days from kindergarten eligibility cutoff date.
Notes: Data are from the restricted-access versions of the 1997–2006 National Health Interview Survey (NHIS). The horizontal axis indicates bins for children born each number of days from the kindergarten eligibility cutoff date. The dots are mean diagnosis rates. The lines are from locally weighted regression interpolation. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff.

In Table 4 we present the regression equivalent to this figure for each data set. The main reduced-form estimates for the NHIS are presented in the top panel of Table 4. The structure of the table mimics that in Table 3 where we start out with a model in column 1 that includes only the coefficient on $I(z_i \geq 0)$, and each successive column adds additional covariates. The estimate reported in column 4 indicates that children born in the 120 days after the cutoff have a 2.1 percentage point lower probability of being diagnosed with ADHD. This corresponds to approximately 24% of the average diagnosis rate across the sample.

Next, results using data from the MEPS are presented in the middle panel of Table 4. We find that being born after the cutoff leads to between a 0.6 and 0.8 percentage point reduction in the probability of being treated for ADHD. This is approximately 13–19% of the mean treatment rate of 4.3% in the sample. Note that in column 4 the results become imprecise and not statistically significantly different from zero once the linear polynomials are included, although the magnitude of the coefficient does not change appreciably.

The bottom panel of Table 4 presents the equivalent set of results using the private claims data base. As in our other data sets, we find that the estimates are not sensitive to the inclusion of demographic characteristics, state and birth cohort fixed effects, or a linear polynomial in days from cutoff. The baseline result in column 4 indicates that children born just after the cutoff experience a 1.6 percentage point lower risk of receiving stimulants to treat ADHD, approximately 27% of the average rate of stimulant usage. The main results reported in Table 4 indicate a large and robust relationship between being born after the kindergarten eligibility cutoff date and being diagnosed with or receiving prescription treatment for ADHD. We find that being born after the cutoff, and therefore being relatively old for grade, is associated with an 13–27% lower risk of ADHD treatment and a 24% lower risk of ADHD diagnosis.

With the caveats about the potential bias in the 2SLS estimates noted above, we can exploit the fact that the model has only one endogenous covariate and the model is exactly identified and combine the estimates from Tables 3 and 4 using Eq. (4) to construct an estimate of the impact of being young for one's grade on ADHD diagnosis and treatment. The 2SLS estimate in the just-identified model can be replicated by dividing the reduced-form estimates in Table 4 by the first-stage estimate, −0.35 (in other words, multi-

[21] Most federal surveys of insurance status tend to under count Medicaid enrollment (Davern et al., 2009). The undercount is large for the NHIS as well with Census Bureau estimates putting the size of the undercount at roughly 25 percent in 2001 and 2002 (http://www.census.gov/did/www/snacc/docs/SNACC_Phase_IV_Full_Report.pdf). Research has suggested that the undercount is primarily due to miscoding the source of the insurance (Call et al., 2008) rather than respondents confusing Medicaid with uninsurance. Lo Sasso and Buchmueller (2004) present evidence that the problem is particularly pronounced for children, suggesting that because of the rise of Medicaid managed care, many with Medicaid report private insurance instead. As a result, we believe a sizeable fraction of people in the NHIS private insurance sample may actually be Medicaid recipients.

Table 4
Regression discontinuity estimates of the effect of being born after the cutoff date.

Covariates	Models			
	(1)	(2)	(3)	(4)
National Health Interview Survey (NHIS)				
Outcome: ADD/ADHD diagnosis				
N = 35,343 children				
Mean of dependent variable = 0. 0864				
Born after cutoff	−0.0204 (.0050)	−0.0209 (.0050)	−0.0206 (.0050)	−0.0208 (.0079)
Age fixed effects, gender, race/ethnicity		×	×	×
State and birth cohort fixed effects			×	×
1st order polynomial				×
Medical Expenditure Panel Survey (MEPS)				
Outcome: receiving medication to treat ADD/ADHD				
N = 31,641 for 18,559 children				
Mean of dependent variable = 0.0427				
Born after cutoff	−0.0055 (.0037)	−0.0059 (.0034)	−0.0063 (.0034)	−0.0079 (.0058)
Age fixed effects, gender, race/ethnicity		×	×	×
State and birth cohort fixed effects			×	×
1st order polynomial				×
Private claims data				
Outcome: prescription claim for Ritalin or other drug for treating ADD/ADHD				
N = 48,206 observations for 22,371 children				
Mean of dependent variable = 0.0584				
Born after cutoff	−0.0124 (.0021)	−0.0123 (.0021)	−0.0122 (.0030)	−0.0156 (0.0057)
Age fixed effects, gender		×	×	×
State and birth cohort fixed effects			×	×
1st order polynomial				×

Notes: Coefficients are from linear probability model regressions with standard errors in parentheses. All specifications include a constant term. All standard errors are clustered by current state of residence. Sample weights are used for the NHIS and MEPS data. The polynomial is defined as days from the cutoff and is modeled separately for days before and days after. The cutoffs are the kindergarten eligibility cutoff date in the child's current state of residence in the year the child turned 5 years old. The variable "Born After Cutoff" is $T(i \geq 0)$. The sample includes children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff date.

plying by 2.85). Therefore our estimates from the NHIS suggest that being young for one's grade increases the chance of being diagnosed with ADHD by 5.9 percentage points (standard error of 2.3 percentage points), or about 70% of the sample mean. From the MEPS, we find that being young for one's class increases the chance of taking stimulants by 2.25 percentage points (standard error of 1.64 percentage points) and the corresponding number for the private insurance sample is 4.45 percentage points (1.62).[22]

The results from Table 4 and the IV estimates in the previous paragraph indicate that relative age is a more important determinant of ADHD diagnosis in comparison to treatment. Given the concerns about stimulant use outlined above, we potentially care more about inappropriate stimulant use than inappropriate diagnosis of ADHD which may go untreated. However, the impact of relative age on stimulant use is a large impact, both in the reduced-form and the IV models.

These results are very similar to the estimates in Elder (this issue). Using data from the ECLS-K survey, Elder demonstrates that a delay in school starting age by 1 year (which by construction would make a student older relative to the median student in a state-year cell) would reduce ADHD diagnosis by 5.4 percentage points and reduce ADHD medication use by 4.4 percentage points. The results are produced by a similar methodology (regression discontinuity design) and model specifications, but with very different samples and over different time periods. Despite these differences the results are remarkably similar across the two studies.

To explore the robustness of the findings reported in Table 4, we perform a variety of specifications checks. It should be noted that for there to be an effect of age relative to the cutoff date on treatment or diagnosis two relationships must be present. First, it must be the case that the kindergarten eligibility laws influence enrollment behavior and therefore age for grade, which is demonstrated in the first-stage regression discussed above. Second, relative age must determine diagnosis and/or treatment for some portion of the population. Given the statistical significance found in Tables 3 and 4, we can infer that both effects are occurring and that there is medically inappropriate diagnosis. It is important to consider heterogeneity in the results to determine whether this average effect is concentrated among a selected or unusual portion of the population. Often in the context of instrumental variable estimation this issue is referred to as determining the Local Average Treatment Effect (LATE), implying that the effect is only measured for individuals that are responsive to the instrument (Angrist et al., 1996).

[22] Treating estimates from all samples as two-sample instrumental variables estimates, one can show that the t-statistic on the reduced form is roughly the t-statistic on the two-sample instrumental variables estimate. Let $t(\hat{\alpha}_2)$ be the t-statistic on the reduced-form (Table 4) and $t(\hat{\gamma}_2)$ be the t-statistic on the first-stage (Table 3). Assuming zero covariance between these two equations, one can show that the squared t-statistic on the 2SLS estimate from Eq. (4) is approximately equal to $t(\hat{\beta}_2)^2 = \{[1/t(\hat{\gamma}_2)]^2 + [1/t(\hat{\alpha}_2)]^2\}^{-1}$. In this case, $t(\hat{\gamma}_2)$ is large in absolute value so $[1/t(\hat{\gamma}_2)]^2$ is very small and close to zero and therefore $|t(\hat{\beta}_2)| \approx |t(\hat{\alpha}_2)|$.

Table 5
Heterogeneity in regression discontinuity estimates of ADHD diagnosis, National Health Interview Survey.

Specification	Sample	Num. of Obs.	Mean of Dep. Var.	Coef. on born after, $T(i \geq 0)$
Baseline results	±120 days	35,343	0.0864	−0.0208 (0.0079)
Days in sample	±90 days	26,659	0.0861	−0.0178 (0.0101)
	±60 days	17,826	0.0849	−0.0203 (0.0136)
	±30 days	9145	0.0826	−0.0316 (0.0155)
Order of polynomial	2nd order	35,343	0.0864	−0.0168 (0.0142)
	3rd order	35,343	0.0864	−0.0381 (0.0181)
	4th order	35,343	0.0864	−0.0364 (0.0205)
	5th order	35,343	0.0864	−0.0523 (0.0255)
State of birth	State of birth	30,476	0.0893	−0.0156 (0.0076)
	State of birth = state of residence	26,607	0.0875	−0.0205 (0.0098)
Gender	Male	18,014	0.1248	−0.0197 (0.0116)
	Female	17,329	0.0471	−0.0209 (0.0088)
Race/Ethnicity	White	19,538	0.1012	−0.0210 (0.0091)
	Black	6000	0.0803	−0.0356 (0.0152)
	Hispanic	8360	0.0462	−0.0220 (0.0134)
Age group	7–12	19,345	0.0818	−0.0237 (0.0107)
	13–17	15,998	0.0926	−0.0158 (0.0114)
Survey years	1997–2001	18,274	0.0798	−0.0180 (0.0096)
	2002–2006	17,069	0.0928	−0.0232 (0.0114)
Privately insured	Private insurance 1997–2006	22,904	0.0807	−0.0259 (0.0102)
	Private insurance, 2003–2006	8240	0.0871	−0.0263 (0.0149)
Additional family-level controls		35,343	0.0864	−0.0209 (0.0076)
Probit model (marginal effects)		35,343	0.0864	−0.0197 (0.0068)

Notes: Data is from the 1997–2006 National Health Interview Survey, and the sample is restricted to children born ages 7–17 on June 1st of the survey year who were born within 120 days of the kindergarten eligibility cutoff. Unless otherwise specified, coefficients are from linear probability regressions with standard errors in parentheses, and all specifications include a constant, a linear polynomial in days from cutoff separately for days before and days after, child's age, state of residence and birth cohort fixed effects, and controls for gender and race/ethnicity. Population weights are used and the standard errors are clustered by state.

In our analysis, we would like to confirm whether the inappropriate diagnosis and treatment we detect is seen across subsets of the population, as well as to confirm whether the empirical results hold with alternative specifications. However, caution must be used in interpreting the relative size of coefficients. It may be the case that some populations are more compliant with the instrument. For example, we know that girls are much less likely to be held back in kindergarten than boys, so are more compliant with the instrument. In that case we might expect to find larger differences across the eligibility cutoff dates, since those dates were more binding for girls than boys.[23] However, it might also be the case that relative age is less important for girls than for boys, due to faster maturation of young girls. In that case, we would expect to see a smaller effect of relative age for girls than for boys. Theoretically, then, it is not obvious whether the coefficient for girls should be smaller or larger than that for boys, or how to interpret any differences between the two. We therefore present these results merely to explore whether the effect holds in subpopulations, but strongly caution against interpreting differences in the coefficients as indicating a stronger or weaker relative age effect. It may simply be that our instrument is more effective at predicting relative age for some populations than others.

Table 5 presents the specification checks for ADHD diagnosis using the NHIS data. Since the coefficient of interest did not change across the columns of Table 4, it is not surprising that in Table 5 we find the estimate is robust to a host of specification checks. These results use the same specification as Table 4, column 4, repeated in the top row of Table 5 for comparison. First, we restrict the window of the sample to children born within 90, 60, and 30 days of the cutoff date. While the estimates become less precise as the sample size decreases, we find that the effect of being born just after the cutoff is an approximately 1.8–3.2 percentage point decrease in the probability of being diagnosed with ADHD. This effect is 21–37% of the total ADHD diagnosis rate. The confidence intervals for each of these estimates overlap meaning that any pair-wise comparison of estimates will not be able to reject the null that the differences across parameters is zero. The results are also insensitive to including higher-order polynomials. These results confirm that the findings cannot be due to season of birth effects.[24]

The third set of sensitivity tests demonstrates that approximating state of residence at age five with state of birth rather than current state of residence produces nearly identical estimates. When we restrict the sample to children that report being born in the same state where they currently reside, a sample much more likely to have been living in that same state at age five, we again find that the results are nearly identical.

Next, we explore the heterogeneity of the estimates across subsets of the population. Note that we include the mean of the dependent variable in the table, which highlights the large differences in diagnosis rates across different groups. We find that nearly 13% of boys have ever been diagnosed with ADHD, compared with 5% of girls. However, we see a similar effect of being born after the cutoff for both boys and girls. This result is not found in our other data sets, where the girls sample does not produce statistically significant effects. In all three datasets we are unable to reject the null hypothesis that the estimates for boys and girls are the same. Although the estimated effect for girls using the NHIS is only slightly larger in magnitude, it is considerably larger in percentage

[23] Indeed, in results not shown, using the full NHIS sample the "first-stage" estimates of the effect of being *born after* on being *Young* are 0.4 for girls compared with 0.3 for boys.

[24] Note that in results not shown but available upon request, estimates are similar when birth month fixed effects are included.

Table 6
Heterogeneity in regression discontinuity estimates of stimulant treatment, Medical Expenditure Panel Survey (MEPS) and private claims samples.

Specification	Sample	MEPS		Private claims	
		Obs. [$\bar{y}$]	Coef. on $T(i \geq 0)$ (Std error)	Obs [$\bar{y}$]	Coef. on $T(i \geq 0)$ (Std error)
Baseline results	±120 days	31,641 [0.0427]	−0.0079 (0.0058)	48,206 [0.0584]	−0.0156 (0.0057)
Days in sample	±90 days	23,744 [0.0410]	−0.0129 (0.0065)	36,582 [0.0572]	−0.0129 (0.0059)
	±60 days	16,034 [0.0391]	−0.0083 (0.0068)	24,809 [0.0563]	−0.0136 (0.0067)
	±30 days	8136 [0.0368]	−0.0104 (0.0125)	12,504 [0.0548]	0.0026 (0.0123)
Order of polynomial	2nd order	31,641 [0.0427]	−0.0143 (0.0070)	48,206 [0.0584]	−0.0064 (0.0059)
	3rd order	31,641 [0.0427]	−0.0033 (0.0111)	48,206 [0.0584]	−0.0103 (0.0132)
Gender	Male	16,109 [0.0610]	−0.0131 (0.0101)	24,216 [0.0803]	−0.0218 (0.0102)
	Female	15,523 [0.0235]	0.0003 (0.0081)	23,990 [0.0363]	−0.0092 (0.0072)
Age group	7–12	18,424 [0.0523]	−0.0039 (0.0086)	23,703 [0.0584]	−0.0150 (0.0080)
	13–17	13,217 [0.0306]	−0.0110 (0.0077)	24,503 [0.0583]	−0.0154 (0.0062)
One observation per claimant	First year in data	16,986 [0.0420]	−0.0064 [0.0064]	19,857 [0.056]	−0.0108 (0.0059)
	Last year in data	14,655 [0.0435]	−0.0102 [0.0068]	19,696 [0.057]	−0.0196 (0.0065)
Probit model (marginal effect)		31,641 [0.0427]	−0.0055 (0.0047)	48,206 [0.0584]	−0.0149 (0.0055)
Private insurance 2003–2006		6570 [0.0481]	−0.0271 [0.0147]		

Notes: Unless otherwise specified, coefficients are from linear probability regressions with standard errors in parentheses. All specifications include a constant, a linear polynomial in days from cutoff separately for days before and days after, child's age fixed effects, and controls for gender, state, year of birth, and, when available, race/ethnicity. In the MEPS population weights are used. In all samples the standard errors are clustered by state. The samples include children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff date.

terms. Note that there may be a power loss when attempting to detect smaller effects on treatment rates, as described below in the discussion of Table 6.

When comparing between different racial/ethnic groups, we find the highest rates of ADHD diagnosis among white non-Hispanic children. While the mean diagnosis rates differ by race, we again find similar coefficients on being born after the cutoff in all samples, with the largest effects for children with Hispanic ethnicity. But again, large standard errors on the lower sized-minority samples mean that any pair-wise comparison is unable to reject the null the coefficients are the same.

So far our estimates have pooled together children ages 7–17. Because a 1 year difference in age represents a larger fraction of a child's life at younger ages, we might expect that the relative age differences cause larger effects for children ages 7–12 compared to teenage children. Note that although all specifications do include child's age and birth cohort fixed effects, we may still find that the rising rates of ADHD diagnosis lead to a larger estimate for the younger age group due to year effects as well.[25] In Table 5, comparing across age groups we find that the largest effect is seen for the youngest age group in the sample.

We then divide the sample into survey years 1997–2001 versus 2002–2006. Consistent with other studies we see that ADHD diagnosis rates rose between these two time periods from 8.0 to 9.3%, or about a 16% rise. The effect of being born after the cutoff is larger in the later time period, 2.3 percentage points (25%) versus 1.8 percentage points (23%) in the earlier time period. This result suggests that the effects of relative age on inappropriate diagnosis may be increasing over time as ADHD diagnosis and treatment become more prevalent.

Next we consider the subset of the population that reports having private health insurance, to approximate a sample that is similar to those in our private claims dataset. These results are reported near the bottom of Table 5. First, note that consistent with previous studies, we find that the diagnosis rates of privately insured children are slightly lower than the national average over this time period due to the higher rates for those on public insurance.[26] Also consistent with previous studies documenting a rise in ADHD diagnosis over time, and with the heterogeneity by survey years described above, we find that when we further restrict the sample to those with private insurance in survey years 2003–2006 we observe higher diagnosis rate. In Table 4, we find a larger impact of being born after the cutoff on ADHD medication use in the private claims sample than in the MEPS sample, which includes those with private and public insurance plus those uninsured. This suggests that children with private health insurance might experience larger effects of being born after the cutoff date on stimulant treatment rates. Similarly, looking at the subset of the population that reports having private health insurance in the NHIS we find that being born after the cutoff reduces the probability of ADHD diagnosis a statistically significant 2.6 percentage points. Although this coefficient is roughly 30% larger than the full sample, the difference in coefficients is not statistically significant. When we reduce the sample to match the survey years of the private claims data there is little change in the estimated coefficient. Therefore, using the NHIS dataset we find an effect for the subset of the population that has private insurance which is slightly larger than that from the population at large, consistent with the disparity found between the MEPS and private claims samples in Table 4.

So far all models used only a limited set of family-level covariates. We have a sparse set of controls available in the private claims sample, and we wanted the models to be as similar as possible. Because covariates do not vary appreciably for those born before and after the cutoff, we do not anticipate that including more detailed family controls will affect the results. To confirm this, we exploit the detailed data in the NHIS. In the next row of Table 5, we add 16 dummy variables for different income values including income not reported, dummy variables for all potential family sizes, dummies for all potential numbers of siblings in the

[25] Note that in results not shown, similar to the findings of Bedard and Dhuey (2006) discussed above, we find that for children age 13–17 the first stage coefficient is only −0.28 compared with a coefficient of −0.41 for the children age 7–12. This is consistent with eligibility being less binding as children age due to differential promotion and retention. It may also be due to using current state of a residence as a proxy for the state where the child lived at age five. As children age it will be more likely that they have moved since age five, so an additional form of measurement error is introduced that may cause attenuation bias.

[26] For a discussion of ADHD diagnosis and treatment rates by health insurance status see, for example, Visser et al. (2007) and references therein.

Table 7
Falsification tests, National Health Interview Survey and private claims data.

Sample	Outcome	Num. of Obs.	Mean of Dep. Var.	Coef. on born after, $T(i \geq 0)$
NHIS	Had chicken pox?	34,727	0.725	−0.0057 (0.0121)
	Have respiratory allergies?	35,233	0.139	0.0059 (0.0086)
	Suffers from hay fever?	35,247	0.127	−0.0046 (0.0089)
	Have frequent headaches?	35,321	0.082	0.0014 (0.0054)
Private claims	Any asthma drug use?	48,026	0.093	0.0117 (0.0075)
	Any antibiotic drug use?	48,026	0.345	0.0080 (0.0091)

Notes: In the top panel data are from the 1997–2006 National Health Interview Surveys and population weights are used. The bottom panel uses the private claims sample. The samples include children ages 7–17 on June 1st of the survey year born within 120 days of the kindergarten eligibility cutoff date. The coefficients are from linear probability regressions with standard errors in parentheses, and the standard errors are clustered by state. All specifications include a constant, a linear polynomial in days from cutoff separately for days before and days after, child's age fixed effects, and controls for gender, state, year of birth, and, in the NHIS only, race/ethnicity.

household and a complete set of dummies for the highest education level in the family. The coefficient on the reduced-form from this model is virtually identical to that in the basic model at the top of the table. Finally, because our dependent variable is dichotomous, we confirm that using a limited dependent variable model produces nearly identical results. The last panel in Table 5 provides the marginal effects from a Probit model; there little impact of changing the estimation method on parameter estimates.

Table 6 reports a similar set of specification and heterogeneity checks considering stimulant prescription as the outcome of interest. In Table 4 we found that the estimated effect of being born after the cutoff was strikingly similar across the specifications as additional covariates were added for all three datasets. The private claims data source has a large enough sample size to explore alternative specifications and heterogeneity within the sample. However in the MEPS data, the main result, presented in column 4 of Table 4, is not statistically significant. Still, we explore whether the qualitative results in the MEPS hold across specifications and within subsamples as further evidence supporting the findings in the larger private claims data. Although it is a smaller data set, the MEPS sample is nationally representative and allows for controls for race and ethnicity.

The top row of Table 6 repeats the main specification, Table 4 column 4, for the MEPS and private claims data sets. In the first set of specification tests, we find that the coefficient is insensitive to narrowing the window of birthdays included in the sample. However in the private claims data, reducing the sample to a 30 day window generates a qualitatively small and statistically insignificant positive coefficient. In the next panel of results in Table 6 we include higher order terms of the polynomials $h(z)$. The specification with the quadratic yields a puzzling result. Using the MEPS data we find the coefficient on *born after* more than doubles, while in the private claims data the coefficient goes to zero. Once higher order terms are added the coefficients are again similar to the baseline result. Recall also that this anomalous result is not found in the NHIS results reported in Table 5. Note that Porter (2003) argues that odd-numbered polynomials have better econometric properties in regression discontinuity design models.

The sensitivity of the results in the private claims data to the window over which we examine the model and the order of the polynomial are in stark contrast to the results from the NHIS which are not sensitive to these model alterations. Upon further inspection, the result can be explained by an anomalously high stimulant use rate on day $z = 6$ (children born 6 days after their state and year-specific cutoff).

There are roughly 200 observations for each day $z = -120$ to 120. On day $z = 6$, the mean stimulant use rate is about 14% which is approximately 2 percentage points higher than any other day and nearly twice the sample average. Looking at the −30 to 30 day models, if we estimate specifications (1) through (3) from Table 4 for this sample, in all cases, the coefficient (standard error) on the treatment effect dummy variable is −0.014 (0.005). However, when we add in the linear terms, the coefficient then drops to the number in Table 6. Estimating the linear time trend on only 30 days when there is an extreme outlier on day $z = 6$ increases the slope on $h(z)$ and eliminates any coefficient on the treatment variable, $I(z \geq 0)$. The distortion due to a large spike in stimulant use on day $z = 6$ is lessened as we increase the window around day $z = 0$.

The outlier on day $z = 6$ also generates havoc with the higher order polynomials since the model is trying to fit the underlying response surface through this one high value for the outcome. If we re-estimate the (−30, 30) day model with linear, quadratic and cubic polynomials in $h(z)$ including a dummy variable for day $z = 6$, then the coefficient on the treatment effect dummy variable in these three models equal −0.0186 (0.0042), −0.0130 (0.0060) and −0.0190 (0.0086), respectively. So although the estimates are sensitive to this one extreme outlier, once the additional dummy variable for day $z = 6$ is included the results are strikingly robust to restricting the sample to smaller windows around the discontinuity and to in the inclusion of higher order polynomials.

We next consider heterogeneity within the samples. As was found with diagnosis rates, treatment rates for boys are much higher than for girls in both samples. The effect of being born after the cutoff is only statistically significant for boys in the private claims data and reflects an over 2 percentage point decreased risk of ADHD treatment for boys born just after the eligibility cutoff. Note that the estimate for girls is not statistically significant, but this may simply be due to insufficient power. Again similar to the estimates in Table 5, we find that the effect is largest for children ages 7–12. Near the bottom of Table 6 we estimate a model in the MEPS data for a subset of the population that is most similar to that from the private claims data base. We find that among children with private health insurance in survey years 2003–2006, 4.8% have a prescription medication to treat ADHD. We find that children born just after the cutoff date have a 2.7 percentage point lower risk of being treated for ADHD in this group.

Table 7 provides our final robustness check. Here we estimate a similar model using other childhood diseases as outcomes, as in Fig. 2. Because children born before the cutoff will have experienced more years of school on average, one might worry that it is exposure to years of school, rather than relative age, that is causing the difference in diagnosis rates. The first two childhood ailments we consider as falsification tests, chicken pox and respiratory allergies, may also be a function of years of exposure to school. Another concern might be that the stress of being younger than one's classroom peers actually causes ADHD. Although we are not aware of any evidence that ADHD is stress-induced, we explore the possibility that children who are relatively young may suffer from stress-induced ailments. To test if a stress-induced mechanism is at work, we consider other childhood ailments that may be exacerbated by stress. For all four childhood ailments we con-

sider, chicken pox, respiratory allergies, hay fever, and frequent headaches, we find no statistically significant effects of relative age. This also further confirms that differences in susceptibility to diseases by children born at different times of year cannot explain the effects.

Similarly, at the bottom of Table 7, if the stress-induced illness or exposure to school mechanism were influencing ADHD treatment, we would expect to see a negative and significant effect of being born after the cutoff on asthma medication use or antibiotic use as well. The estimates at the bottom of Table 7 show a positive and statistically insignificant effect of relative age on asthma medication and antibiotic use.

In summary, the estimates of the effect of being born after the age of school start date are large and statistically significant across a host of specifications and in almost all subsamples. We find no similar effect for several other childhood diseases and conditions. This suggests that the nature of the diagnostic guidelines, which recommend a comparison with classroom peers, leads to medically inappropriate ADHD diagnosis.

As discussed above, one concern is that our results are potentially driven by the fact that children that are young for their grade will have also been in school 1 year longer. We are worried that "exposure" to diagnosis in schools could lead to higher diagnosis rates, even controlling for age and birth cohort. This does not appear to explain our results. First, note that in Table 6 when we restrict the sample to include only children 13–17 years we still produce a statistically significant reduction in ADHD treatment rates of 1.5 percentage points for those born just after the cutoff. As children age, the relative difference in exposure declines considerably and hence, if exposure were driving the differences in outcomes across groups, we should find little impact in this older group. Second, recall that the sensitivity tests presented in Fig. 2 and in Table 7 indicated that other childhood diseases and conditions that might be influenced by differential exposure do not show a similar pattern. Third, we next consider the sensitivity of our estimates to including grade fixed effects to the model to explicitly control for years of exposure. However, we do not present these results in Table 5 because this model is problematic for two reasons. As discussed in detail above, we note that the grade in school variable is measured with considerable error. But more importantly, it is also the case that years of schooling is an outcome, so is endogenous. Therefore we view these results solely as a sensitivity test to compare children in the same grade, which confirms the coefficient is not being driven by exposure. Since years of schooling and $I(z_i \geq 0)$ are negatively related, it is not surprising that adding grade fixed effects decreases the coefficient on $I(z_i \geq 0)$ to -3.8 percentage points (standard error of 0.9 percentage points). Hence we can be reassured that even when comparing within grade by including grade fixed effects we find a large and statistically significant effect of being born after the cutoff. These three sets of results indicate that exposure to years of schooling cannot explain the findings.

In summary, if one assumes that the true incidence rate of ADHD is uniform over a small window around the age at school start cutoff, the estimates provide compelling evidence that a large fraction of ADHD diagnoses are not the result of an underlying medical condition. Rather, children that were born just after the kindergarten eligibility cutoff date in their state in the year they turned 5 years old, who therefore were more likely to wait an additional year to enter school, are at a much lower risk being diagnosed with ADHD and being prescribed stimulants. This provides strong evidence that medically inappropriate diagnosis and treatment is occurring.

The diagnosis rates for children born on either side of the kindergarten eligibility cutoff date should only be different if that cutoff date actually corresponds to initial school enrollment behavior. Not only do many states allow exemptions for early entry, in general states do not require children attend school until they are 7 years old. In addition, more advanced children may skip grades, while children who are struggling may repeat grades. This noncompliance with the age at school start laws should only serve to dampen the difference between children born before and after the cutoff date. As described above, we cannot estimate the effect of being relatively young directly due to data limitations. Still, the reduced-form analysis presented here indicates that, as long as the underlying medical risk of having ADHD does not differ across the eligibility cutoff date, there is a significant amount of medically inappropriate diagnosis and treatment of ADHD.

6. Conclusions

The evidence presented above indicates that for some children, a diagnosis of ADHD is not solely based upon underlying biological conditions. Rather, being born just before versus just after the kindergarten eligibility cutoff date in one's state is a significant factor in the probability of receiving an ADHD diagnosis. This is likely a result of relative maturity and is therefore not a surprise given the difficulty of diagnosing ADHD and the explicit consideration that health care providers are advised to give to whether the behaviors in question "happen more often in this child compared with the child's peers?"[27] As Elder and Lubotsky (2009) demonstrate, younger children in classes are more likely to have educational and behavioral problems compared to their peers, and therefore, some children who are relatively young compared to their classroom peers are more likely to be diagnosed with ADHD. These results suggest that the comparison sample for diagnosis should not be other children in class but rather, other children of a similar age within a class.

Note that even if it is the case that children entering school at younger ages triggers ADHD, this would suggest an important causal mechanism that the medical research should further explore. ADHD is now thought to stem from both neurological and environmental factors. If being exposed to formal schooling at younger ages is actually causing a rise in ADHD, we must then revisit educational policy and consider how children are segmented into classrooms and how age-appropriate educational activities are chosen.

Our econometric model does not, however, allow us to identify whether particular children who are young for their grade are overdiagnosed or whether some older children are under-diagnosed. It could be the case that younger children are over-diagnosed because they are acting immaturely relative to classroom peers (but not relative to children the same age), and this behavior is misinterpreted as indicating the child has ADHD. Alternatively, it could also be that because of the stark age difference between children born before and after the cutoff in early grades, it is easier to diagnose younger children with ADHD and older children are left under-diagnosed. Indeed, as described in the beginning of Section II, under-diagnosis of ADHD and other mental health disorders is a significant and important public health concern as children with ADHD are at an increased risk of academic difficulties and are more likely to engage in risky behaviors. Children with ADHD often have difficulties in school and untreated ADHD may lead to lower human capital accumulation, although research suggests that many children taking stimulant medication may suffer from toxicity that will hamper rather than improve their cognitive function (Swanson et al., 1991). In addition, recent research has highlighted the externalities associated with having a child with extreme emotional or

[27] NIMH, ADHD Booklet, Page 6.

behavioral problems in the classroom, such as the lower human capital accumulation of classmates (Aizer, 2009; Fletcher, 2010). What our results do indicate is that observationally similar students have very different diagnosis rates depending on when their birthday falls in relation to the start of the school year. As such, the results suggest that for a large fraction of children their current medical diagnosis with regard to ADHD is not based on underlying biological factors.

That said, we do feel that evidence from outside our econometric model is suggestive that the more likely scenario is that younger children are inappropriately diagnosed as having ADHD when they are in fact simply less mature than their peers. Evidence from brain imaging technology suggests that ADHD is associated with a 3 year developmental delay in a child's brain (see, e.g., Shaw et al., 2010). The NIMH ADHD Booklet describes children with ADHD as being hyperactive, inattentive, and/or impulsive. Inattentiveness could be missed in older children, but extreme hyperactivity and impulsivity are unlikely to go unnoticed. And while it is theoretically possible that older children would have symptoms that would not be detected by teachers, that notion is not consistent with the idea that children with ADHD have severe and uncontrollable behavioral problems. Along similar lines, children's whose ADHD symptoms are not severe enough to be detected may not require medical intervention and treatment. Thus, even if our results suggest that over- and under-diagnosis are both occurring, the possibility of over-diagnosis due to relative immaturity may be of more significant public health concern than under-diagnosis in this scenario.

Potential over-diagnosis of stimulant medication is particularly troubling given the possible side effects of these drugs. According to a 2007 FDA review, the stimulant medications used to treat ADHD have rare but serious and significant potential side effects including cardiovascular problems and psychiatric problems.[28] Others studies have suggested potential long-term consequences on young children's brain development. According to our estimates, approximately 9% of all children are diagnosed with ADHD and approximately 4–6% of children current take a prescription stimulant to treat ADHD. According the population estimates provided by the U.S. Census Bureau,[29] on July 1, 2006 there were approximately 53 million children ages 5–17 in the United States. To put our estimates into perspective, an excess of 2 percentage points implies that approximately 1.1 million children received an inappropriate diagnosis and over 800,000 received stimulant medication due only to relative maturity. Recognizing the pattern of inappropriate diagnosis should help to better target treatments. In addition, this may help to avoid treatments with potentially serious short-term and long-term consequences.

International comparisons that indicate the United States spends more yet achieve lower health outcomes when compared to other OECD countries. This and other evidence has prompted criticism of wasteful spending and over-treatment in the U.S. healthcare system. However, identifying inappropriate diagnosis and treatment can be difficult and generally involves costly chart reviews or extensive case studies. In this paper we document inappropriate medical diagnosis and treatment using survey data. Using variation in relative age induced by age of school start laws, we are able to clearly identify a source of differential diagnosis that cannot be due to true underlying differences in disease incidence.

[28] Findings from the FDA review are described in the NIMH ADHD Booklet, Page 9.

[29] Source: Population Division, U.S. Census Bureau, Table 2: Annual estimates of the resident population by sex and selected age groups for the United States: April 1, 2000 to July 1, 2008 (NC-EST2008-02), Release Date: May 14, 2009, accessed November 16, 2009.

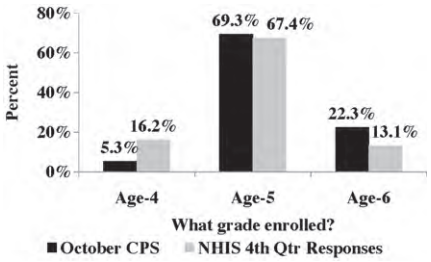

Fig. A1. Current grade for children 7–16, 2000–2002 October CPS and 4th quarter responses to 2000–2002 NHIS.
Notes: The NHIS fourth quarter responses are from the public use data. We impute the respondents' ages as of October 1st using information on month and year of birth. Both samples are from years 2000 to 2002 for children age 7–16.

Appendix A.

Fig. A1.

References

Ahmann, P.A., Waltonen, S.J., Theye, F.W., Olson, K.A., Van Erem, A.J., 1993. Placebo controlled evaluation of Ritalin side effects. Pediatrics 91 (6), 1101–1106.
Aizer, A., 2009. Peer effects, institutions and human capital accumulation: the externalities of ADD. Working paper April 2009.
Angold, A., Erkanli, A., Egger, H., Costello, J., 2000. Stimulant treatment for children: a community perspective. Journal of the American Academy of Child and Adolescent Psychiatry 39 (8), 975–984.
Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91 (434), 444–455.
Angrist, J.D., Krueger, A.B., 1992. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. Journal of American Statistical Association 87 (418), 328–336.
Baicker K., Chandra A., 2004. Medicare spending, the physician workforce and beneficiaries' quality of care. Health Affairs Web Exclusive, W4-184. http://content.healthaffairs.org/cgi/content/abstract/hlthaff.w4.184.
Barkley, R., Guvermont, D., Anastopoulos, A., DuPaul, G., Shelton, T., 1993. Driving-related risks and outcomes of attention deficit hyperactivity disorder in adolescents and young adults: a 3- to 5-year follow-up survey. Pediatrics 92 (2), 212–218.
Bedard, K., Dhuey, E., 2006. The persistence of early childhood maturity: international evidence of long-run age effects. The Quarterly Journal of Economics 121 (4), 1437–1472.
Biederman, J., Wilens, T.E., Mick, E., Faranone, S.V., Spencer, T., 1998. Does attention-deficit hyperactivity impact the developmental course of drug and alcohol abuse and dependence. Biological Psychiatry 44 (4), 269–273.
Byrd, R.S, Weitzman, M., Auinger, P., 1997. Increased behavioral problems associated with delayed school entry and delayed school progress. Pediatrics 100 (4), 654–661.
Call, K.T., Davidson, G., Davern, M., Nyman, R., 2008. Medicaid undercount and bias to estimates of uninsurance: new estimates and existing evidence. Health Services Research 43 (3), 901–914.
Castle, L., Aubert, R.E., Verbrugge, R.R., Khalid, M., Epstein, R.S., 2007. Trends in medication treatment for ADHD. Journal of Attention Disorders 10 (4), 335–341.
Cox, E.R., Motheral, B.R., Henderson, R.R., Mager, D., 2003. Geographic variation in the prevalence of stimulant medication use among children 5 to 14 years old: results from a commercially insured U.S. sample. Pediatrics 111 (2), 237–243.
Currie, J., Stabile, M., 2006. Child mental health and human capital accumulation: the case of ADHD. Journal of Health Economics 25 (6), 1094–1118.
Datar, A., 2006. Does delaying kindergarten entrance give children a head start? Economics of Education Review 25 (1), 43–62.
Davern, M., Klerman, J.A., Baugh, D.K., Call, K.T., Greenberg, G.D., 2009. An examination of the Medicaid undercount in the Current Population Survey: preliminary results from record linking. Health Services Research 44 (3), 965–987.
Dhuey, E., Lipscomb, S., forthcoming. Disabled or young? Relative age and special education diagnoses in schools. Economics of Education Review.
Dobkin, C., Ferreira, F., 2010. Do school entry laws affect educational attainment and labor market outcomes? Economics of Education Review 29 (1), 40–54.
Elder, T.E., this issue. The importance of relative standards in ADHD diagnoses: evidence based on exact birth dates. Journal of Health Economics.
Elder, T.E., Lubotsky, D., 2009. Kindergarten entrance age and children's achievement: impacts of state policies, family background, and peers. Journal of Human Resources 44 (3), 641–683.
Fisher, E.S, Wennberg, D.E., Stukel, T.A., et al., 2003. The implications of regional variation in Medicare spending. Part 2: Health outcomes and satisfaction with care. Annals of Internal Medicine 138 (4), 288–298.

Fertig, M., Kluve, J., 2005. The effect of age at school entry on educational attainment in Germany. IZA Working Paper No. 1507.
Fletcher, J.M., 2010. Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. Journal of Policy Analysis and Management 29 (1), 69–83.
Fuchs, V., 2004. More variations in use of care, more flat-of-the-curve medicine. Health Affairs, Web Exclusive, October 7, 2004. http://content.healthaffairs.org/cgi/content/full/hlthaff.var.104/DC2.
Institute of Medicine, 2007. Learning what works best: the nation's need for evidence on comparative effectiveness in health care. http://www.iom.edu/ebm-effectiveness.
Goodman, R., Gledhill, J., Ford, T., 2003. Child psychiatric disorder and relative age within school year: cross sectional survey of large population sample. British Medical Journal 327 (743), 472–475.
Jensen, P.S., Martin, D., Cantwell, D.P., 1997. Comorbidity in ADHD: implications for research, practice and DSM-V. Journal of the American Academy of Child and Adolescent Psychiatry 36 (8), 1065–1079.
LeFever, G.B, Dawson, K.V., Morrow, A.L., 1999. The extent of drug therapy for attention deficit-hyperactivity disorder among children in public schools. American Journal of Public Health 89 (9), 1359–1364.
Lincove, J.A., Painter, G., 2006. Does the age that children start kindergarten matter? evidence of long-term educational and social outcomes. Educational Evaluation and Policy Analysis 28 (2), 153–179.
Lo Sasso, A.T., Buchmueller, T.C., 2004. The effect of the state children's health insurance program on health insurance coverage. Journal of Health Economics 23 (5), 1059–1082.
McEwan, P.J., Shapiro, J.S., 2008. The benefits of delayed primary school enrollment: discontinuity estimates using exact birth dates. Journal of Human Resources 43 (1), 1–29.
Mayes, S.D., Calhoun, S.L., Crowell, E.W., 2000. Learning disabilities and ADHD: overlapping spectrum disorders. Journal of Learning Disabilities 33 (5), 417–424.
Morrill, M.S., 2008. Women's labor supply and the family. University of Maryland, PhD Dissertation.
MTA Cooperative Group, 1999. A 14-month randomized clinical trial of treatment strategies for attention-deficit hyperactivity disorder (ADHD). Archives of General Psychiatry 56 (12), 1073–1086.
National Institute of Mental Health, 2008. Attention deficit hyperactivity disorder. Bethesda (MD): National Institute of Mental Health, National Institutes of Health, U.S. Department of Health and Human Services. NIH Publication No. 08-3572. [Revised 2008; Accessed November 2009]. http://www.nimh.nih.gov/health/publications/attention-deficit-hyperactivity-disorder/adhd_booklet.pdf.
Nissen, S.E., 2006. ADHD drugs and cardiovascular risk. The New England Journal of Medicine 354 (14), 1445–1448.
Pelham, W.E., Foster, E.M., Robb, J.A., 2007. The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. Ambulatory Pediatrics 7 (18), 121–131.
Pliszka, S.R., 1998. Comorbidity of attention-deficit/hyperactivity disorder with psychiatric disorder: a overview. Journal of Clinical Psychiatry 59 (Suppl. 7), 50–58.
Porter, J., 2003. Estimation in the regression discontinuity model. Working paper, Department of Economics, University of Wisconsin.
Puhani, P.A., Weber, A.M., 2007. Does the early bird catch the worm? Instrumental variable estimates of early educational effects of age of school entry in Germany. In: Dustman, C., Fitzenberger, B., Machin, S. (Eds.), Studies in Empirical Economics: The Economics and Training of Education. Physica–Verlag, Heidelberg, Germany, pp. 105–132.
Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., Sobek, M., 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. University of Minnesota, Minneapolis.
Safer, D.J., Malever, M., 2000. Stimulant treatment in Maryland public schools. Pediatrics 106 (3), 533–539.
Sax, L., Kautz, K.K., 2003. Who first suggests the diagnosis of attention-deficit/hyperactivity disorder. Annals of Family Medicine 3 (1), 171–174.
Shaw, P., Gogtay, N., Rapoport, J., 2010. Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. Human Brain Mapping 31, 917–925.
Stipek, D., Byler, P., 2001. Academic achievement and social behaviors associated with age of entry into kindergarten. Journal of Applied Developmental Psychology 22 (2), 175–189.
Swanson, J.M., Cantwell, D., Lerner, M., McBurnett, K., Hanna, G., 1991. Effects of stimulant medication on learning in children with ADHD. Learning Disabilities 24 (4), 219–230.
Trampush, J.W., Miller, C.J., Newcorn, J.H., Halperin, J.M., 2009. The impact of childhood ADHD on dropping out of high school in urban adolescents/young adults. Journal of Attention Disorders 13 (2), 127–136.
Visser, S.N., Lesesne, C.A., Perou, R., 2007. National estimates and factors associated with medication treatment for childhood attention-deficit/hyperactivity disorder. Pediatrics 119 (Suppl. 1), S99–S109.
Volkow, N., Insel, T., 2003. What are the long term effects of methylphenidate treatment? Biological Psychiatry 54 (12), 1307–1309.
Wennberg, J.E., Fisher, E.S., Goodman, D.C., Skinner, J.S., 2008. Tracking the Care of Patients with Severe Chronic Illness – the Dartmouth Atlas of Health Care. Dartmouth Institute for Health Policy and Clinical Practice, Center for Health Policy Research, Lebanon, NH.
West, J., Meek, A., Hurst, D., 2000. Children Who Enter Kindergarten Late or Repeat Kindergarten: their Characteristics and Later School Performance (NCES No. 2000-039). U.S. Department of Education, Washington, DC.
Woodward, L.J., Fergusson, D.M., Horwood, L.J., 2000. Inattentive and hyperactive behaviors and driving offenses in adolescence. American Academy of Child and Adolescent Psychiatry 39 (5), 627–634.
Zito, J.M., Safer, D.J., dosReis, S., Gardner, J.F., Boyles, M., Lynch, F., 2000. Trends in the prescribing of psychotropic medications to preschoolers. Journal of the American Medical Association 283 (8), 1025–1030.
Zuvekas, S.H., Vitiello, B., Norquist, G.S., 2006. Recent trends in stimulant medication use among U.S. children. American Journal of Psychiatry 163 (4), 579–585.