I Z A

Consequences of ADHD Medication Use for Children's Outcomes

Søren Dalsgaard
Helena Skyt Nielsen
Marianne Simonsen

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Consequences of ADHD Medication Use for Children's Outcomes

Søren Dalsgaard
Aarhus University

Helena Skyt Nielsen
Aarhus University
and IZA

Marianne Simonsen
Aarhus University
and IZA

ABSTRACT

Consequences of ADHD Medication Use for Children's Outcomes[*]

This paper estimates effects of early ADHD medication use on key human capital outcomes for children diagnosed with ADHD while using rarely available register based data on diagnoses and prescription drug purchases. Our main identification strategy exploits plausible exogenous assignment of children to hospitals with specialist physicians, while our analysis of health outcomes also allows for an individual level panel data strategy. We find that the behavior of specialist physicians varies considerably across hospitals and that the prescribing behavior does affect the probability that a given child is treated. Results show that children diagnosed with ADHD in pharmacological treatment have fewer hospital contacts if treated and that treatment to some extent protects against criminal behavior.

JEL Classification: I11, I12

Keywords: Attention Deficit/Hyperactivity Disorder, treatment, hospital variation, health, crime

Corresponding author:

Marianne Simonsen
Aarhus University
Fuglesangs Alle 4
DK-8210 Aarhus V
Denmark
E-mail: msimonsen@econ.au.dk

"TO date, no study has found any long-term benefit of attention-deficit medication on academic performance, peer relationships or behavior problems, the very things we would most want to improve. Until recently, most studies of these drugs had not been properly randomized, and some of them had other methodological flaws."

- Professor Emeritus of Psychology, L. Alan Sroufe, Minnosota: "Ritalin Gone Wrong", New York Times, January 12, 2012.

I. Introduction

One of the most publicly debated new medical technologies is pharmacological treatment of children diagnosed with Attention-Deficit/Hyperactivity-Disorder (henceforth ADHD). There are concerns, even among some professionals in the field, that children are over-diagnosed with ADHD; that we know too little about effects of ADHD medication and that children are, in fact, over-treated with ADHD medication. This paper is one of the first to investigate longer-run effects of pharmacological treatment of ADHD on key human capital variables such as health and criminal behavior. Particularly little is known about the latter type of outcome. More generally, our paper speaks to the literatures on the impacts of early health interventions and the importance of non-cognitive skills and mental health.

ADHD is one of the most common chronic mental health problems among young children. It is estimated to affect about 3-7 % of all children (The Danish Association for Child and Adolescent Psychiatry (2008); American Psychiatric Association (2000)) or on average one child – more often a boy than a girl – in every classroom. The underlying prevalence is found to be constant across many countries and cultures, including North America, Australia, South America, and Europe; see Faraone et al (2003). Core symptoms associated with ADHD are attention deficiencies, hyperactivity and impulsiveness and children often simultaneously suffer from other behavioral problems along with depression and anxiety. Hence, ADHD is likely to affect not only one's overall human capital but also one's tendency to engage in risky health behaviors.

Children with ADHD grow up in relatively disadvantaged families and their ADHD may even contribute to family vulnerability (Kvist et al. (2013)). They have – even when in pharmacological treatment – much worse long-term outcomes than others who are comparable in terms of age and gender (Dalsgaard et al. (2002); Mannuzza and Klein (2000)) and also when compared to siblings without ADHD (Currie and Stabile (2006); Fletcher and Wolfe (2008)). And although a series of

randomized controlled studies show that treatment with central nervous system stimulants[1] is effective in terms of reducing ADHD core symptoms and improving social behavior (van der Oord et al. (2008)), most studies only have very short follow up periods of up to three years after randomization implying that we effectively know very little about the longer-run consequences of treating children pharmacologically. Because these studies rely on relatively small sample sizes and since attrition has shown to be a real concern in the follow-up studies already carried out (see van der Oord et al (2008)), alternative evaluation methods are called for.

We propose to exploit variation in access to pharmacological treatment generated by hospital variation in specialist physicians' propensities to prescribe to estimate effects of treatment. We use a combination of Danish registers that apart from rich socio-economic background variables include the following key information: 1) psychiatric history and diagnoses for children and their parents including information about the hospital where a diagnosis was first established, 2) history of prescription drug usage for children and their parents, and 3) measures of health and crime.

We find that the behavior of specialist physicians varies considerably across hospitals and that the prescribing behavior does affect the probability that a given child is treated. Results show that children diagnosed with ADHD on the margin of receiving pharmacological treatment have *fewer* hospital contacts if treated and have *fewer* interactions with the police.

In the analysis of health outcomes, we have sufficient data to meaningfully distinguish between birth cohorts. We document that effects are smaller in later cohorts where more children are diagnosed and treated pharmacologically before the age of ten. There are still significant gains from treatment in the later cohorts, but the results support a hypothesis of diminishing returns to broadening the group of treated.

The paper is structured as follows: Section II discusses causes of ADHD and links between ADHD and human capital, Section III presents the empirical framework, and Section IV the institutional background for the analysis. Section V shows the data, Section VI the results and Section VII concludes.

---

[1] Methylphenidate is the most common pharmacological treatment for ADHD, better known under the brand name Ritalin. A more recent development is Concerta; a once daily extended release formulation of methylphenidate. Another commonly used agent is dextroamphetamine, however no extended-release formulations of this are available in Denmark and the use therefore very limited.

II. Causes of ADHD and Links between ADHD and Human Capital

Recently, a series of papers such as Cunha et al. (2006), Currie (2011), and Currie and Almond (2010) have emphasized the importance of investing early in particularly vulnerable children. Moreover, Cunha and Heckman (2007) show theoretically that early investments not only have a large potential pay-off, they are also efficient in the sense that an equity-efficiency trade-off does not exist, which is the case for later investments. The reasons are that skills acquired in one period persist into future periods and that skills produced at one stage raise the productivity of investment at subsequent stages. Importantly, skills are multidimensional and are likely to complement each other. The group of children with ADHD is a prime example for which we would expect early investments with immediate effects on health capital in general to also have long-term consequences for later health and human capital attainment. This paper investigates investments via pharmacological treatment. Early take-up of pharmacological treatment may have long-term effects on human capital simply because it improves behavior and therefore the likelihood of future treatment but also because of dynamic complementarities: treatment may improve cognitive skills including less impulsive behavior and more awareness of the consequences of one's actions that feed back on outcomes.

When modeling effects of pharmacological treatment, it is clearly critical to be aware of the causes of ADHD since they may be correlated with initial conditions that determine later outcomes. Though not perfectly described, it is well known that genetic factors are very important (Faraone and Doyle (2001)), but also premature birth, birth complications, maternal smoking and alcohol use during pregnancy are associated with ADHD (Linnet et al. (2003)). In our sensitivity section we therefore investigate whether effects of treatment vary with health at birth and maternal smoking during pregnancy. Children with ADHD are also more likely than others to have language, cognitive and memory problems (e.g. Jensen et al. (2001) and Frazier et al. (2004)). To address this, we investigate the extent to which results are affected by children with mental retardation (11% in our sample). Previous literature has shown that ADHD worsens learning outcomes; see for example Currie and Stabile (2006), Fletcher and Wolfe (2008), and Ding et al (2009).

As mentioned above, it is well-documented that treatment with central nervous system stimulants is effective in terms of reducing the number and impact of ADHD core symptoms.[2] This evidence is based on a series of randomized controlled trials. In the seminal Multimodal Treatment Study of Children with ADHD (henceforth MTA), 579 children aged 7-9.9 years suffering from ADHD were assigned to different types of treatment for a period of 14 months. Of these, 144 children were assigned to pure pharmacological treatment. Within the 14-month period careful medication management with or without behavioral treatment was shown to be superior to routine community care or behavioral treatment in terms of reducing core symptoms, see MTA (1999). In follow-up studies considering children three years after randomization, the difference in symptom relief diminishes over time and eventually disappears (see MTA (2004) and Molina et al. (2009)). The studies find no impact of medication management on functioning outcomes such as social skills, relations or reading achievement. While being informative about symptom relief, these studies cannot stand alone when it comes to determining the long-term consequences of pharmacological treatment of ADHD. Unfortunately, the MTA study, as well as other randomized controlled trials, suffers from problems such as crossover and contamination after intervention; see also Molina et al. (2009).

There is mixed evidence with regards to the long-term consequences of pharmacological treatment of ADHD on human capital accumulation based on observational methods. Paykina and Greenhill (2008) report less school disruption, anti-social behavior and academic failure following pharmacological treatment, while others raise doubt about such effects (e.g. Mendez et al. (2011)). A recent paper by Currie, Stabile and Jones (2013) exploits a reform of health care insurance in Quebec, Canada that lowered prices on all prescription drugs to investigate effects of Ritalin use on emotional functioning and academic outcomes for children with parent reported types of behavior associated with ADHD. They find that Ritalin use significantly worsened emotional functioning for girls and academic outcomes for boys. Regarding accumulation of health capital, we are aware of only one study by Marcus et al. (2008) that considers the link between pharmacological treatment and health. Their study uses a duration model to investigate the association between compliance in pharmacological treatment and injuries for a group of children in treatment. They find that children treated with high intensity had a non-significantly lower risk of injury than those treated with low

---

[2] The information is surveyed and incorporated in national treatment guidelines e.g. The Danish Association for Child and Adolescent Psychiatry (2008) for Denmark and Paykina and Greenhill (2008) for the US.

intensity. Regarding crime outcomes, a recent study by Lichtenstein et al (2012) uses Swedish register-based data to document the association between treatment receipt in adulthood and crime for individuals with ADHD while exploiting that some individuals periodically drop out of treatment.

III. Empirical Framework

III.A *IV strategy*

The key problem when evaluating the effects of ADHD and its treatment on human capital formation is how to identify relevant counterfactuals. For every child in pharmacological treatment for ADHD, we would ideally like to know his or her outcome in the absence of treatment. Since this is never observed, we instead ask the question: how do we find a non-treated candidate who is similar in terms of observable and unobservable characteristics except for the fact the he is not being treated? In order to credibly identify causal effects, we exploit plausible exogenous variation in access to pharmacological treatment generated by hospital variation in psychiatrists' propensity to prescribe: imagine two children, both diagnosed with ADHD and with the same characteristics; one will be treated but only because he meets a psychiatrist with preferences for using pharmacological treatment while the other does not. While we only have access to hospital level information about physicians' behavior, this strategy is inspired by Duggan (2005) who uses the same type of variation to investigate effects of second-generation antipsychotics on spending on other types of medical care, Doyle (2007, 2008) who uses variation in investigator assignment to estimate causal effects of foster care, and Maestas, Mullen, and Strand (2013) who rely on random examiner variation to look at the effects of disability insurance on labor market attachment. Such variation may stem from hospital level differences in treatment culture and knowledge spillovers; see Coleman, Katz and Menzel (1957) and Soumerai et al. (1998). We investigate hospital level treatment cultures in detail in Section VI.A below.

The instrument is based on filled prescriptions, and therefore it reflects *both* the probability that the physician write a prescription and the probability that the patient fills a prescription. Both the former (aggressiveness in prescribing behavior) and the latter (better hospital level engagement with families) is potentially relevant hospital level variation in our case. The filling probability is likely

correlated with socio-economic characteristics (predictive of treatment motivation or engagement) which we control for in our analysis.

Let Y indicate a human capital outcome such as criminal behavior. X is a set of observable characteristics that determines both the propensity to receive pharmacological treatment and the outcome. PH is an indicator for receiving pharmacological treatment for ADHD.

Formally, we model the relationship between child outcomes and treatment receipt:

(1) $$Y^i = X^i\beta + \alpha^i PH + \varepsilon^i$$

or alternatively

(2) $$Y^i = X^i\beta + \bar{\alpha}PH^i + \left(PH^i(\alpha^i - \bar{\alpha}) + \varepsilon^i\right),$$

where the term in the squared brackets is the error term. Clearly, PH may be correlated with ε if, for example, unobserved parental investments leads to an increased likelihood of pharmacological treatment. Also, PH may be correlated with α if pharmacological treatment is based on expected (and foreseeable) gains. To solve this, we implement an instrumental variables strategy that exploits hospital level differences in propensities to prescribe.

Observable and unobservable child and parental characteristics (X and θ) may affect whether the child with ADHD receives pharmacological treatment ($PH = 1$), yet the prescribing specialist physicians are instrumental in making the decision; see details below. Let Z be a hospital level measure of the physician's propensity to prescribe. We model treatment receipt as:

(3) $$PH^i = 1\left[Z^i\gamma + X^i\beta + \theta^i > 0\right].$$

For the instrumental variables strategy to be valid, we need the hospital level propensity to prescribe to positively affect the likelihood that a given child diagnosed with ADHD receives pharmacological treatment. This is clearly testable (formally, we test the null $\gamma = 0$). A second, un-testable, identifying assumption is that physicians do not affect children's outcomes beyond their choice of treatment. Therefore, it is crucial that individuals are not selective in their choice of hospitals and that physicians who are more likely to prescribe pharmacological treatment to children with ADHD do not systematically use other types of treatment too; see discussion below. A final assumption (already imposed in (3) via the common coefficient γ) is monotonicity: if child i receives pharmacological treatment when met with a physician with a low propensity to prescribe,

7

then child *i* must also receive treatment if met with a physician with a high propensity to prescribe. With our continuous instrument, we identify an Average Marginal Treatment Effect (AMTE) rather than the usual LATE associated with binary instruments. This parameter is the average treatment effect for children on the margin of receiving pharmacological treatment.

III.B *Individual level fixed effects panel data strategy*

In practice treatment is rarely initiated before the age of five and never before the age of four.[3] For health outcomes, where we have access to individual level panel data, we therefore also employ a difference-in-differences strategy, comparing outcomes for treated children prior to (age 4) and after treatment (in our main analysis age 10 +) with untreated diagnosed children before and after, corresponding to a fixed effects or first difference analysis; see Blundell and Costa Dias (2009) and Lechner (2011). With heterogeneous treatment effects, this will provide estimates of the average treatment effect on the treated (ATET).

---

[3] We exclude one child treated before the age of five.

FIGURE 1

PROBABILITY OF AT LEAST ONE HOSPITAL VISIT AT GIVEN AGE,

BY TREATMENT STATUS[*]
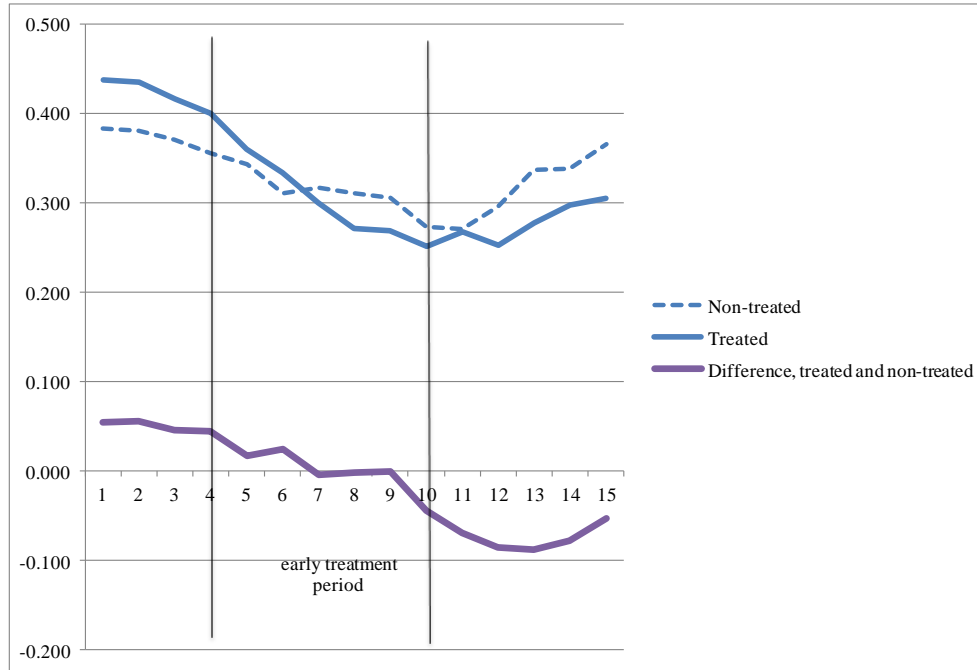
Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Details about the measurement of the outcome and the sample follow below.

This identification strategy allows for selection into treatment based on, for instance, severity of symptoms, hospital and parental characteristics *as long as* these influences are constant over time. Thus, even if particularly attentive parents are systematically more (or less) likely to engage in pharmacological treatment *and* more (or less) likely to use health care services at any time, it will not violate the identifying assumptions. However, if attentive parents are more (or less) likely to engage in pharmacological treatment *but only* more (or less) likely to use health care services when children are below 5, this would indeed violate the identifying assumptions.

Our main regression model for the difference-in-differences strategy is the following:

$$Y_{ia} = \alpha_i + \beta_1 1(age > 9)_i + \beta_2 treat_i \cdot 1(age > 9)_i + \varepsilon_{it}$$

where Y is the outcome of interest, *treat* indicates that the child belongs to the treatment group (i.e. receives pharmacological treatment before age ten), $1(age > 9)$ indicates post-treatment age, ε is an error term, i indexes individuals, and a indexes age, β_2 is the parameter of interest. In practice, all

background variables are measured prior to or at childbirth and thus do not vary across time. The effect of these variables will therefore be cancelled out along with the individual level fixed effect, α_i.

The key identifying assumption in a difference-in-differences set-up is that there can be no differential trends between the treatment and control group in the absence of treatment. Figure 1 illustrates the relevant identifying variation and main results for the probability of hospital visits in a given year. The figure demonstrates that treated children have a higher probability of interacting with general hospitals prior to diagnosis and treatment than non-treated children, but that (except for the year of birth when children are very rarely exposed to injury) the development is otherwise parallel. After the diagnosis is established and treatment is initiated), this tendency is reversed, however. After the age of seven, treated children – who were initially more disadvantaged – perform better than non-treated children.

IV. Institutional Set-up

This section describes the decision stages and agents involved in diagnosing and treating ADHD. We consider three stages: The first step involves the seeking of a referral for evaluation at the specialist-level, the second step the establishment of a diagnosis, and the final step the treatment decision.

IV.A *Seeking of a diagnosis and hospital assignment*

Parents – and if not parents then in some cases school psychologists, teachers, or school nurses – decide whether to seek a referral for evaluation in the first place. Typically, this involves a visit to the family's general practitioner (GP) who serves as a gatekeeper for specialist treatment. The GP can then – if he agrees with the indications – provide parents with a referral to a specialist at a child and adolescent psychiatric hospital. Unfortunately, child and adolescent outpatient clinics experience considerable waiting lists to first contact of on average 70-80 days.[4] For the purpose of our instrumental variables strategy, however, this is convenient because it conceivably weakens the degree to which parents are able to control the choice of clinic. To a large extent parents will have to accept the available hospital for their child and to the best of our knowledge they will have no

---

[4] See the Danish Health and Medicines Board, www.sst.dk

systematic information about the treatment patterns of their assigned hospital prior the first visit.[5] We investigate the correlation between the type of clinic and a long list of child and parental background characteristics and also perform a range of robustness tests intended to shed light on the importance of clinic selection. In the vast majority of cases, specialist physicians are child and adolescent psychiatrists, but pediatricians and neurologists also do assessments and diagnose. There is substantial variation in the size of the specialist clinics at the hospitals; some employ as few as two physicians while some have as many as 28.[6] Consultations with the GP are free of charge (for the parents) as are those with specialist physicians when equipped with a referral from the GP.

IV.B *Diagnoses*

Diagnoses are based on the *International Classification of Diseases* (ICD) diagnostic manual developed by WHO.[7] Diagnosis and treatment practices are documented by the Danish Association for Child and Adolescent Psychiatry (2008). Measurements of psychopathology such as the Child Behavior Checklist (CBCL) have been standardized in Danish (Bilenberg, 1999) and have been part of the standard clinical assessment in most child and adolescent psychiatric clinics in Denmark since the early 1990s. Multi-informants are always used in the assessment of children at hospital-based child and adolescent psychiatric units and a standard assessment often includes a direct observation of the child by a trained psychiatric nurse at the day-care/school and at home with the family is often part of the assessment and also a test of the cognitive level by a psychologist. In contrast to the US (Evans et al. (2010); Elder (2010)) there is no Danish evidence that being relative-young-for-grade increases the probability of receiving an ADHD diagnosis (Dalsgaard et al.

---

[5] If all regional general hospitals are overbooked, the GP may refer the child to a private practicing specialist. It *is possible* to consult with a specialist at a private clinic without a GP reference, but the parents must then pay the costs themselves.

[6] Information gathered from the Danish Medical Association.

[7] ADHD is classified as an F90 diagnosis. This covers hyperkinetic disorders, activity and attention disorders, other hyperkinetic behavioral disorders, and hyperkinetic behavioral disorders without further specification. See WHO (1993). F90 actually corresponds to one of the three subtypes in DSM, namely the combined type. In ICD-10 the diagnosis for the inattentive subtype ADHD is F98.8, but the validity of this diagnosis in the registers has not been examined and ICD has no manualized or operationalized criteria for this disorder, as is does for F90.

(2012); Dalsgaard et al. (2013)). This would be problematic because it suggests that the ADHD diagnosis is not entirely objective and that relative standards are being applied.[8]

IV.C *Pharmacological treatment*

Given an ADHD diagnosis, the specialist may recommend pharmacological treatment. This typically implies treatment with Methylphenidate (98% of the treated children in our estimation sample). Methylphenidate is almost exclusively used to treat ADHD symptoms.[9] Medications used in the treatment of ADHD all act to increase brain catecholamine level. Although Methylphenidate has been used therapeutically for more than 60 years, the precise prefrontal cortical and subcortical mechanisms of action are poorly understood, but are associated with its ability to block the dopamine and norepinephrine re-uptake transporters (Solanto (1998)). It is well-known, however, that dopamine increases attention, interest and motivation. Common side effects are insomnia, headaches, decreased appetite, increased blood pressure and heart rate, and symptoms of depression and anxiety. Prescription drugs are heavily (and universally) subsidized by the government. The maximum yearly out-of pocket costs for individuals with chronic diseases like ADHD are $ 670.[10] In fact, most children with ADHD will, according to the Law of Service, § 41, be eligible for a subsidy that covers the full costs of prescription drugs.

Parents may, of course, refuse pharmacological treatment. Thus, both the specialist a child meets, the severity of early symptoms, and parental preferences may impact on the likelihood of being treated. Regardless of the choice of pharmacological treatment and the severity of the condition, the reference program advises that children with an ADHD diagnosis are offered social skills training. This is typically the responsibility of the schools and not determined by the clinic at which the diagnosis was obtained.

---

[8] Clearly, this does not prove that diagnoses in Denmark are entirely objective. As discussed below in Section V and shown in Table A1, children with an ADHD diagnosis are negatively selected in terms of socioeconomic background. Whether this represents some type of discrimination is difficult to say.

[9] It may, however, also be used to treat the rare condition of narcolepsy.

[10] 2013 maximum limit and exchange rate on October 10, 2013.

V. Data

Our starting point is the population of Danish children born in the period from 1990-1999. The main data stem from the Danish Psychiatric Central Register; see Munk-Jørgensen and Mortensen (1997) for a detailed description. These data include information about psychiatric history and diagnoses for parents and children diagnosed with Danish general hospitals. The data cover the period from 1960-2010 for the adult population but before 1994, information about children's psychiatric diagnoses was not available. Because it is extremely rare that children are diagnosed before age 4, we include children born as early as 1990.

The psychiatric registers are not constructed for research, but for administrative purposes. Diagnoses are therefore clinical diagnoses, not the result of a systematic well-described uniform psychiatric assessment. The validity of the diagnoses of ADHD in the Danish Psychiatric Central Register has previously been shown to be good, however. The agreement percentage on a full diagnosis of ADHD according to the American Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) was 89%, while the remaining 11% lacked only 1 symptom to fulfill the ADHD diagnosis (Linnet et al. (2009)).

Via unique personal identifiers, the information from the Danish Psychiatric Central Register is merged with registers containing rich socio-economic background variables (from 1980-2007), in-patient somatic disease histories (from 1980-2010), and prescription drug purchases including data for both parents and children (from 1997-2010). Our crime data stem from two sources: a) the official crime register that records charges from the age of 15 (the age of criminal responsibility) up until 2011 and b) national police records of all interactions regardless of the age of the involved parties up until 2012.

TABLE 1

OBSERVABLE CHARACTERISTICS AT CHILDBIRTH, 4,556 CHILDREN BORN 1990-1999

WITH AN ADHD DIAGNOSIS ESTABLISHED BEFORE THE AGE OF TEN

Variable	Treatment before age 10 (N=1456)		No treatment before age 10 (N=3100)	
	Mean	Std. Dev.	Mean	Std. Dev.
Child:				
Boy (0/1)	0.86	0.35	**0.83**	0.37
5-minute APGAR score	9.78	0.86	**9.71**	1.20
Birthweight less than 1,500 grams (0/1)	0.02	0.12	0.02	0.14
Birthweight, 1,500-2,500 grams (0/1)	0.06	0.23	0.07	0.26
Birthweight, above 2,500 grams (0/1)	0.93	0.3	**0.91**	0.29
Complications at birth (0/1)	0.32	0.47	**0.27**	0.44
Gestation length (weeks)	39.28	2.53	39.17	2.53
Mental retardation (0/1)	0.12	0.33	0.11	0.31
Mother:				
Age at childbirth	27.93	4.85	28.24	5.28
High school or less (0/1)	0.50	0.50	0.53	0.50
Length of education (years)	11.40	2.31	11.41	2.30
Unemployed less than 13 weeks (0/1)	0.82	0.38	**0.79**	0.41
Unemployed 13-26 weeks (0/1)	0.12	0.33	**0.14**	0.35
Unemployed more than 26 weeks (0/1)	0.06	0.24	0.07	0.25
Employed in November (0/1)	0.54	0.50	**0.50**	0.50
Gross income (DKK, 2004 prices)	176823	69947	172928	76761
Psychiatric diagnosis (0/1)	0.10	0.30	0.10	0.30
Heart disease (0/1)	0.03	0.18	0.03	0.17
Respiratory disease (0/1)	0.21	0.41	**0.18**	0.38
Smoker (0/1)	0.24	0.43	**0.16**	0.36
Father:				
Age at child birth	30.99	6.06	31.34	6.34
High school or less (0/1)	0.44	0.50	0.44	0.50
Length of education (years)	11.29	2.31	11.34	2.39
Unemployed less than 13 weeks (0/1)	0.86	0.35	**0.84**	0.37
Unemployed 13-26 weeks (0/1)	0.08	0.27	0.08	0.27
Unemployed more than 26 weeks (0/1)	0.06	0.24	**0.09**	0.28
Employed in November (0/1)	0.86	0.35	0.84	0.37
Gross income (DKK, 2004 prices)	264250	128305	256665	152014
Psychiatric diagnosis (0/1)	0.08	0.27	0.08	0.26
Heart disease (0/1)	0.04	0.20	0.04	0.19
Respiratory disease (0/1)	0.14	0.35	0.13	0.34

Notes: Bold indicates that mean for treated children is significantly different from the mean for non-treated children at the 5 % level. With the exception of mental retardation, which is diagnosed in connection with the ADHD diagnosis, all variables are measured in the year just prior to the birth of the child or in connection with childbirth. One child is excluded because of treatment before age 5.

We define *early pharmacological treatment* as purchases in an amount that corresponds to at least six months of treatment in a given year before the age of ten.[11] We think of our analyses as measuring the effects of early take-up of treatment. Pharmacological treatment of ADHD consists of Amphetamine (N06BA01), Methylphenidate (N06BA04), and Atomoxetine (N06BA09).[12]

To assure that early treatment is relevant for all children in our sample, we select from the original data of children born 1990-1999 the 4,557 children who have been diagnosed with ADHD at Danish general hospitals before the age of ten. These children and their parents are clearly disadvantaged in terms of background characteristics as compared to their non-ADHD counterparts: children suffering from ADHD have worse birth outcomes, their parents have lower levels of education, are more likely to be unemployed and have lower income, are more likely to have a psychiatric diagnosis themselves and have a higher prevalence of both heart disease and respiratory disease and mothers are much more likely to smoke during pregnancy; see Appendix Table A1. Among children with an early diagnosis, we delete one child who was treated before the age of five to make sure that our complementary individual level panel data analysis is not contaminated. This gives us a final sample of 4,556 children.

Table 1 presents means of background characteristics by early pharmacological treatment status. Though some differences in background variables are statistically significant, it is not clear that treated children are either more advantaged or disadvantaged than non-treated children. One exception is that mothers of children in treatment are far more likely to have smoked during pregnancy than mothers of non-treated children and also to have respiratory diseases.[13] Among the advantageous characteristics, children in treatment have slightly higher 5-minute APGAR scores and are more likely to have a birthweight of 2,500 grams or more. Their parents are also slightly more likely to be employed.

Our sample corresponds to about 0.6 % of the children in the relevant cohorts. The share of a cohort diagnosed with ADHD before the age of ten has been increasing from 0.3 % in the 1990 cohort to

---

[11] This corresponds to 182 defined daily doses (30 mg Methylphenidate) in a calendar year.

[12] Of course, it would be interesting to evaluate the effects of the duration of treatment too; unfortunately, this is complicated by the fact that the group of children who continues in treatment is also more likely to consist of those who experience a gain.

[13] See Obel et al. (2011), who use sibling differences to show that if smoking is a causal factor behind hyperkinetic disorders, it only has a minor impact.

1.2 % in the 1999 cohort and boys are almost four times more likely than girls to receive an ADHD diagnosis. Similarly, among children diagnosed early, the share in early treatment increases across cohorts from 19 % in the 1990 cohort to 43 % in the 1999 cohort. The share of children ever diagnosed with ADHD in Denmark (about 2 %) is below the predicted prevalence of 3-7 %. Of course, some of this is caused by the fact that we only include children diagnosed at general hospitals. Still it suggests that children in our sample are relatively severe cases.

Treated children may, of course, continue in treatment and also non-treated children may receive treatment later in life. This is important to be aware of when interpreting our formal results below. Figure 2 shows the share of children receiving any pharmacological treatment for ADHD at age ten or later by their early treatment status. We see that most children in early treatment continue after the age of nine but the share is declining with age. Children who are treated for less than six months before the age of ten (our non-treated children) also receive some pharmacological treatment at age ten and later but are much less likely to do so. Again, the share decreases with age. Strictly speaking, therefore, we measure the effects of early treatment initiation *including* the effects of continuing.

While there were no large differences in background characteristics for the groups of treated and non-treated children, as seen in Figure 1 above treated children do exhibit more risky health behavior prior to treatment than non-treated children in that they have more contacts with general hospitals. As shown by Dalsgaard et al (forthcoming), treated children also have more early injuries than non-treated children. We interpret this type of early health behavior as a proxy for the severity of ADHD in a child and we exploit this directly in our sensitivity analyses. Of course, it will also be important to control for early symptoms at the individual level.

FIGURE 2

PROBABILITY OF RECEIVING TREATMENT FOR AT LEAST 6 MONTHS
AT A GIVEN AGE, 4,556 CHILDREN BORN 1990-1999 WITH AN ADHD DIAGNOSIS
ESTABLISHED BEFORE THE AGE OF TEN.
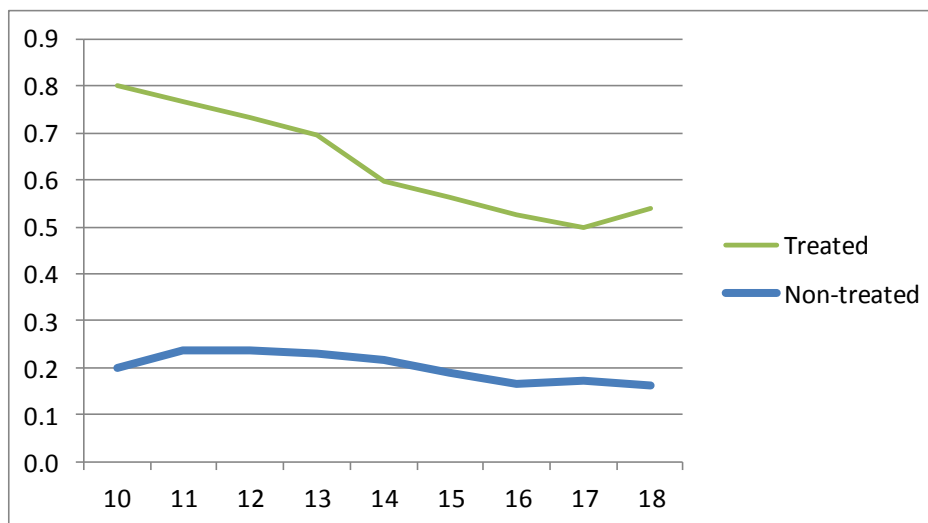
Notes: Treatment status is defined as pharmacological treatment for at least six months in a year before the age of ten. One child excluded because of treatment before age 5.

Our estimations will consider two types of outcomes: 1) health outcomes (one or more contacts with general hospitals at age 10 (excluding hospital visits associated with the treatment of ADHD), number of contacts with general hospitals at age 10 (again excluding hospital visits associated with the treatment of ADHD); one or more contacts with the emergency ward at age 10; number of contacts with the emergency ward at age 10) and 2) crime outcomes (one or more interactions and number of interactions with the police at ages 12-15; at least one charge at ages 16 and 17). For the health outcomes that are measured at age 10, we will rely on all birth cohorts from 1990-1999 while analyses of crime outcomes will use only subsets of these cohorts. Information about charges at age 16, for example, is available until 2011 allowing us to use the 1990-1995 cohorts. Outcome means are shown with the results tables.

VI. Results

As discussed above, our main estimation strategy uses hospital level variation in propensities to prescribe pharmacological treatment to identify causal effects of treatment. This section first details

the instrument and discusses hospital assignment. After this, we present our analysis of the effects of pharmacological treatment.

VI.A *Hospital variation and assignment*

Formally, we define our instrument as: for child *i* the share of treated among *other* children born in the same cohort and diagnosed at the same hospital. Thus the focal child is removed before the share is calculated. To be included in our IV analysis, the hospital must diagnose at least one other child born in the same cohort.[14] In the *health sample* we rely on variation in the treatment propensity among 422 hospital-cohort combinations with on average just above nine children per hospital-cohort combination. Importantly, our analysis below includes regional[15] and cohort fixed effects. Therefore, we essentially exploit variation in hospital level behavior (towards children born at the same point in time) relative to the behavior of physicians at other hospitals in the broader region. Figure 3 shows a stylized example of the geographic set-up: a given region will have a limited number of hospitals and children from within the region will have a chance to go to either of them depending on who has an available specialist, independently of the general practitioner the child sees or the school and police district children belong to. Importantly, school and police districts will be much smaller than the typical catchment areas of the hospitals. To the extent that more (or less) successful school districts, for example, are systematically found in regions where psychiatrics prescribe more, this should be captured by our region fixed effects.

In this section, we conservatively present the results for the instruments while relying on the larger *health sample*. Our conclusions do not change when considering the smaller estimation samples; in fact because of the smaller sample size, the set of tests that we apply to our instruments to render probable that the identifying assumptions are not violated obviously have less power with fewer observations.

---

[14] Results are robust to excluding hospitals that diagnose less than 5/10 other children and to conditioning on treating at least one other child born in the same cohort. This is important to the extent that our measure of the tendency to prescribe is less precise in hospitals that diagnose and treat fewer children.

[15] In our main analysis, regional dummies indicate diagnosis region.

FIGURE 3

STYLIZED GEOGRAPHIC SET-UP
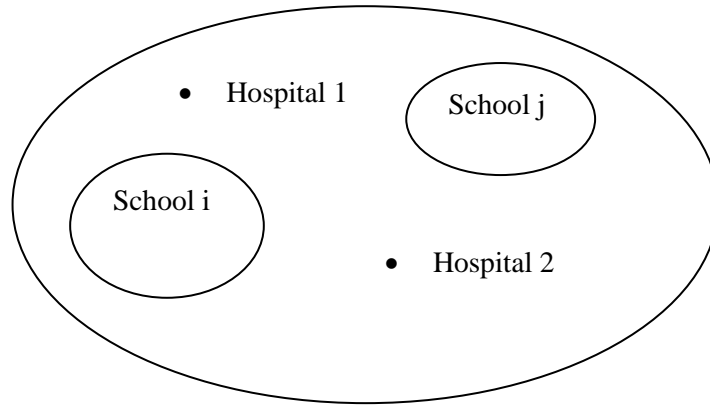
FIGURE 4

DISTRIBUTION OF HOSPITAL LEVEL PROPENSITIES TO PRESCRIBE
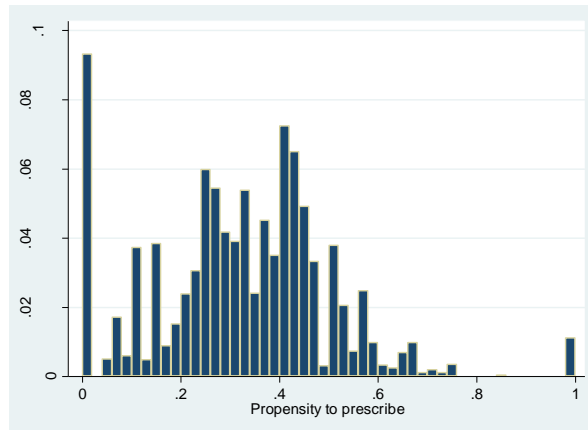
Notes: The propensity to prescribe is defined as the share of other children born in the same cohort and diagnosed at the same hospital.

Figure 4 shows the distributions of the propensity to prescribe. Clearly, there is very little support above values of 0.6. In practice, we will rely on variation in the instrument that lies strictly within the area of support and the results should be interpreted accordingly.

To gain insights into the nature of this considerable variation, we graph hospital specific treatment propensities across cohorts. While there is, as expected, some noise in these processes, we are able to categorize hospitals into four classes: a) consistent high treaters, b) trend followers, c) consistent low treaters, and d) regime jumpers. Figure 5 shows an example of each type.

19

FIGURE 5

COHORT VARIATION IN HOSPITALS' PROPENSITIES TO TREAT
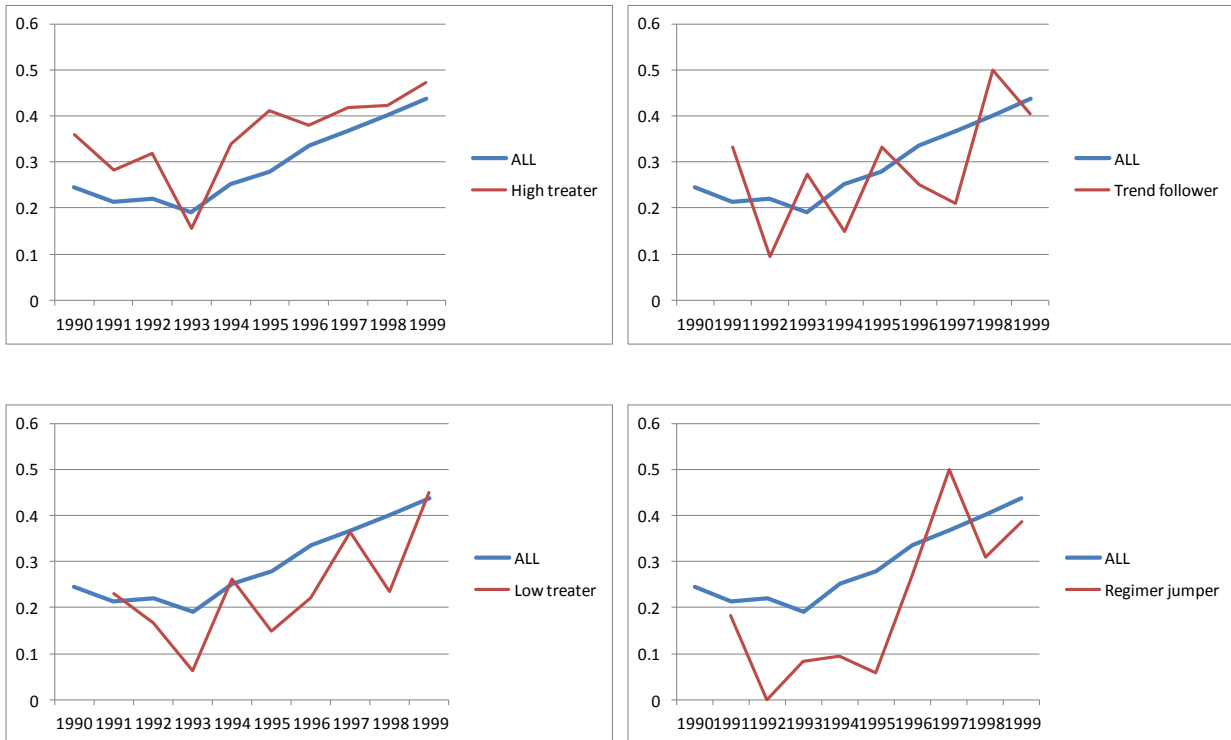
Notes: Time variation in treatment propensities, selected hospitals.

Adjusting for the overall trend, hospitals in the first three categories – the vast majority – exhibit fairly stable treatment behaviors across cohorts of diagnosed children. This supports our hypothesis that differences in treatment cultures drive observed differences in treatment propensities. We do see some hospitals change their behavior dramatically over time; this always reflects an upwards jump in treatment propensity and sometimes happens after a period with low overall activity (fewer than 10 diagnosed children per cohort). Interestingly, we see that "*regime jumpers*" are most often smaller psychiatric units that diagnose relatively few children or non-psychiatric departments where children are typically diagnosed by neurologists or pediatricians; see Section IV.A. As long as *regime jumps* are seen for exogenous reasons such as change of management or the arrival of new information about use of treatment elsewhere, our identification strategy will still be valid. Jumps in the severity of ADHD of the patients diagnosed at the department, on the other hand, will be problematic for our strategy. Department (and hence physician) type also contributes to the variation in the probability of early treatment. As expected, psychiatric departments are significantly more likely to use pharmacological treatment: 35% of children diagnosed at a

psychiatric department are treated before the age of 10 compared to 26 % of children diagnosed at other types of hospital departments.

One issue with our strategy is parental agency in pursuing treatment. The concern is that parents of children with more severe ADHD will be more aggressive in securing appointments with physicians who are more likely to prescribe stimulant medications. This will cause a downwards bias in our results. Or we might see that more resourceful and skilled parents seek diagnosis and treatment, which might cause an upward bias in our results. We first test how our instrument relates to background characteristics of the child and his family (and simple hospital characteristics: i) whether the child is diagnosed at a hospital that receives more than the median number of children and ii) a proxy for whether the hospital only sees "mild cases", defined as children with less than the median number of injuries before age 5) in a simple OLS regression. We find that although a few coefficients are statistically significant, most are small in size. Importantly, we find very low correlations between the instrument and early injuries, parental education and labor market attachment. Results are available on request. Another way to think about this issue is as selective clinic choice: it is possible that some parents move to get access to pharmacological treatment via a hospital with a higher propensity to prescribe if they believe strongly in a positive effect. This would be problematic for our estimation strategy because it would imply a potential correlation between the instrument and gains from treatment. First remember that our instrument is largely uncorrelated with parental and child characteristics suggesting that this issue is of minor importance. Second, in our robustness analysis below we show that a) constructing our instrument at the county level instead in order to minimize local strategic choices and b) basing our instrument on place of birth instead of diagnosis do not change our results.

A closely related issue with our strategy is so-called treatment by indication: physicians at hospitals who are observed to treat more just do so because they meet children who suffer from more severe ADHD. An indirect test of this phenomenon is to consider the relationship between the instrument and the propensity to diagnose at the population level: if living in an area where hospitals treat more heavily also implies that a child from the overall population is, for example more (or less) likely to be diagnosed in the first place, we would worry that high treatment hospitals see different types of children. We estimate a probit for diagnosis receipt and condition on the full set of background variables in addition to the propensity to prescribe at the hospitals within the county (i.e. the hospitals a child living within the county would have a risk of meeting) and conclude that there is

little relationship between the risk of receiving a diagnosis and the original instrument. The results are available on request and imply that if the propensity to treat in county increases by 10 percentage points (relative to a mean of 0.32), this will increase the number of diagnoses relative to the mean with 0.00014/0.006= 2.3 %.

Another issue is whether hospitals that treat more early also differ in terms of other practices. Since social skills training is primarily the responsibility of the schools, we are mostly concerned about other medical treatment provided by the specialist physicians. Of course, if the amount of social skills training and pedagogical back-up in schools is lowered as a consequence of medical treatment, our estimates will likely be biased. However, in this case (and assuming that such additional interventions improve child outcomes), our strategy most likely estimates a *lower bound* on the effects of medical treatment. For 1998-1999 birth cohorts we have information about early enrolment into schools designed for children with special needs. We find no differences in the likelihood that treated and non-treated children receive such additional pedagogical treatment; the share of treated children enrolled at age 8 (9) is 0.193 (0.204) whereas the share of untreated children enrolled at age 8 (9) is 0.191 (0.218). Finally, we saw above that children who are treated early are also more likely to receive treatment later in their life. We find, however, no strong association between the instrument and use of the five most common other nervous system drugs.[16] In the same way, we also investigate whether our instruments are associated with age at diagnosis and find no evidence in favor of this hypothesis: an increase in age at diagnosis of one year is associated with a reduction in the overall propensity to prescribe of 0.002. Strictly speaking, therefore, our instruments measure the propensity to prescribe more early and to continue longer but it does not capture other factors such as special schools, other types of medication, or diagnosis age. This is the treatment regime to have in mind when interpreting our results.

---

[16] Results are robust to excluding children who are treated with these five other types of nervous system drugs (Valproic Acid, Lamotrigine, Risperidone, Clopenthixol, and Diazepam).

TABLE 2

FIRST STAGE RESULTS, HEALTH SAMPLE

	Coef. Est.	Std. Err.	Coef. Est.	Std. Err.	Coef. Est.	Std. Err.
Propensity to treat	**0.485**	0.054	**0.294**	0.067	**0.329**	0.067
Hospital variables:						
Big hospital (0/1)					0.011	0.014
A hospital treating only mild cases of ADHD (0/1)					0.016	0.029
Child variables:						
Injury and poisoning* (DS00-DT98), age 0 (0/1)					-0.028	0.045
Injury and poisoning* (DS00-DT98), age 1 (0/1)					**0.047**	0.021
Injury and poisoning* (DS00-DT98), age 2 (0/1)					0.000	0.018
Injury and poisoning* (DS00-DT98), age 3 (0/1)					0.007	0.017
Injury and poisoning* (DS00-DT98), age 4 (0/1)					*0.033*	0.018
Boy (0/1)					**0.053**	0.019
5-minute APGAR score					-0.013	0.059
Birth weight less than 1,500 grams (0/1)					**-0.128**	0.061
Birth weight, 1,500-2,500 grams (0/1)					-0.060	0.033
Complications at birth (0/1)					0.009	0.017
Gestation length (weeks)					-0.006	0.004
Mental retardation diagnosis (0/1)					0.034	0.022
Mother variables:						
Age at child birth					-0.003	0.002
Length of education (years)					-0.002	0.004
Unemployed less than 13 weeks (0/1)					-0.003	0.032
Unemployed 13-26 weeks (0/1)					-0.028	0.036
Employed in November (0/1)					0.025	0.021
Gross income, kr. 100.000 (2004 prices)					0.008	0.013
Psychiatric diagnosis (0/1)					-0.002	0.022
Heart disease (0/1)					0.034	0.037
Respiratory disease (0/1)					0.029	0.017
Smoker (0/1)					*0.041*	0.023
Father variables:						
Age at child birth					-0.002	0.001
Length of education (years)					-0.002	0.004
Unemployed less than 13 weeks (0/1)					0.019	0.029
Unemployed 13-26 weeks (0/1)					0.015	0.038
Employed in November (0/1)					0.020	0.026
Gross income, kr. 100.000 (2004 prices)					0.007	0.006
Psychiatric diagnosis (0/1)					0.013	0.027
Heart disease (0/1)					0.017	0.041
Respiratory disease (0/1)					0.005	0.024
Cohort and regional dummies	NO		YES		YES	
Sample size			4287			
R-squared	0.027		0.039		0.052	

Notes: Bold (italic) indicates that coefficient is significant at the 5 (10) % level. Instrument lies in 0-0.6 interval. Hospitals included if they diagnose at least two children born in the same cohort. A big hospital receives more than the median number of children. Mild cases are children with less than the median number of injuries before age 5. Standard errors clustered at cohortXcounty level. One child excluded because of treatment before age 5.

VI.B *First stage results*

Table 2 shows the first stage results for the *health sample*. We see that the relationship between the probability of treatment receipt and the instrument is positive: being exposed to a high-intensity treatment regime increases the probability that a given child is treated. The first column shows results without additional covariates while the second column adds cohort and regional control variables and the third column adds child and parental characteristics. As expected given the time variation in prescription rates, adding cohort and regional control variables reduces the relationship between treatment receipt and our instrument but further conditioning on background characteristics leaves the coefficient estimate associated with the instrument unchanged. Figure 6 documents that the relationship between the probability of treatment receipt (actual and predicted) and the instrument is in fact monotonous over the support of the instruments. As argued by Doyle (2007), this provides some support for the monotonicity assumption discussed above.

FIGURE 6

HOSPITAL LEVEL PROPENSITIES TO PRESCRIBE AND THE

PROBABILITY OF EARLY TREATMENT
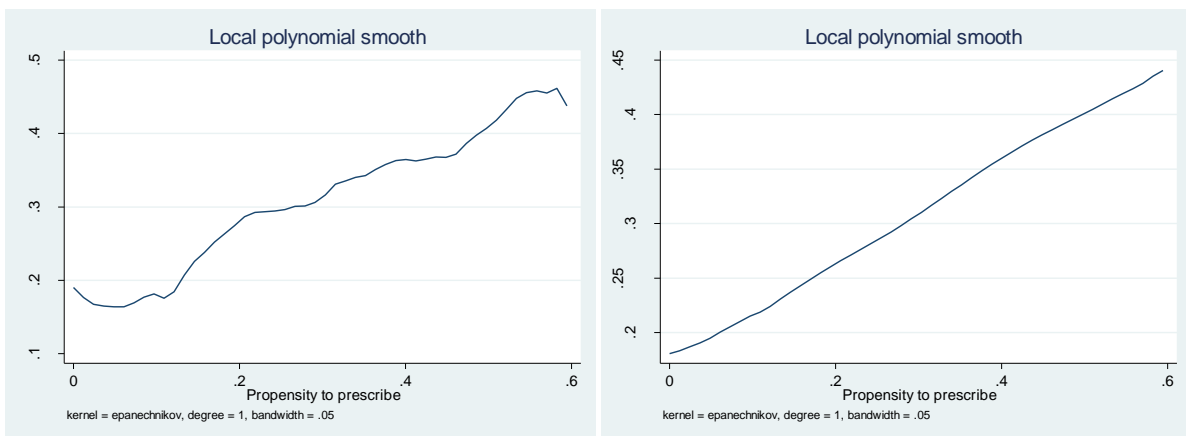
Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Left graph shows results without control variables, right panel results with control variables (same as Table 2). One child excluded because of treatment before age 5.

VI.C *Effects of treatment on health and crime outcomes: OLS and IV Results*

OLS and IV results are shown in Table 3. We first investigate the effects of treatment on health outcomes. We show results with cohort and regional dummies included and results that include the entire conditioning set. Including individual level covariates merely adds precision to the estimates.

We see that treatment is effective in reducing interactions with general hospitals at age 10, both in terms of the incidence and number of hospital visits. Additional analyses, available on request, indicate that this is driven by a reduction in injuries. In line with this, we also find some reductions in emergency ward visits. The reason that we see more precisely estimated effects on injuries than on emergency ward visits is that injuries may also be treated outside of the emergency ward. The OLS results are much smaller in size than the IV results; to the extent that treated children are negatively selected this is expected.[17] Remember also that the IV analysis estimates average effects for children on the margin of receiving pharmacological treatment and this parameter is rarely directly comparable to the OLS. Still, since estimates in some cases exceed the mean, it has to be the case that those affected by treatment are those with a high baseline level risk of injuries. The large coefficients do suggest caution in interpretation however.

We next investigate crime outcomes. The OLS results suggest no benefits associated with treatment. This pattern is reversed in the IV analyses; here we find large and significant gains from treatment at ages 12-13. Parameter estimates are similar at older ages though, possibly because of smaller samples, primarily insignificant. Crime results for ages 14-16 are shown in Appendix Table A2.

Threats to validity in the IV strategy

We discussed above that our instrument measures propensities to prescribe more early and to continue longer but it is informative about the ADHD medication tendency alone. We also saw that there is little tendency for families to move to counties where hospitals treat more intensively. We formally address issues of treatment by indication and strategic settlement in this section. Due to space considerations, we restrict ourselves to the main health outcomes. The full set of results is available on request. First, we use an alternative instrument based on the propensity to treat children with less severe ADHD, namely the share of treated among other children *with weak symptoms* as proxied by a low number of pre-treatment injuries and poisoning. This is precisely an attempt to hold fixed the severity of ADHD. We define the number of early injuries and poisoning to be low if the sum of these before the age of five lies below the median. Of course, this may primarily capture hyperactivity/impulsiveness or inattentiveness in children, which may be rather select consequences

---

[17] In principle, our IV strategy allows for estimation of marginal treatment effects. Unfortunately, our sample is too small for us to conclude much from these.

of early ADHD. Results are shown in column 3 in Table A3. Second, we address potential selection of hospitals within a county by defining our instrument at the county level instead of at the hospital level. In this analysis, county is defined at the time of diagnosis (column 4). Finally, we investigate strategic settlement in the time between the birth of the child and the actual diagnosis by using variation in treatment propensities in the birth county (column 5). Results vary a little between the main specification and the three additional robustness analyses but the overall conclusion is the same.

VI.D *Effects of treatment on health outcomes: Individual level fixed effects panel data results*

As discussed above, our individual level panel data information on health outcomes allows for a difference-in-difference analysis with individual level fixed effects. Remember that this analysis compares outcomes for treated children prior to (age 4) and after treatment (age 10) with untreated but diagnosed children before and after; see Figure 1. Though we saw above that hospitals that are more likely to prescribe do not tend to use other medication more, we cannot completely rule out that such hospitals are simply better in other but *unobserved* dimensions. Such hospital specific effects will be out in a panel data analysis. Table 4 shows the formal results. Again, due to space considerations, we conservatively show results for the indicator for interactions with general hospitals and emergency wards but other results are available on request. For comparison, we reproduce simply the OLS results in columns 1-2. In line with the IV analysis, we find that early pharmacological treatment is effective in significantly improving health outcomes: the probability of being in contact with a general hospital at least once is reduced with 7 percentage points, corresponding to just below 30% of the mean at age 10.

The panel data analysis actually provides for a test of the importance of hospital specific effects beyond differences in their use of prescription drugs: Table A4 shows panel data estimates of the effects of treatment for hospitals with below and above the median value of the instrument. Importantly, we find no significant differences between these two sets of results.

TABLE 3

EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE AGE 10

		OLS		IV	
	Mean		(2)	(3)	(4)
Health outcomes, age 10:					
First stage				**0.294**	**0.329**
				0.067	0.066
# Hospital contacts, age 10	0.498	**-0.069**	**-0.074**	**-1.169**	**-0.800**
		0.033	0.033	0.542	0.441
Hospital contacts, age 10 (0/1)	0.278	*-0.025*	*-0.027*	-0.220	-0.166
		0.014	0.014	0.153	0.156
# Emergency ward visits, age 10	0.233	-0.025	*-0.030*	-0.241	-0.280
		0.018	0.018	0.224	0.203
Emergency ward visits, age 10 (0/1)	0.181	-0.010	-0.014	-0.057	-0.109
		0.013	0.013	0.149	0.131
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		4556			4287
Crime outcomes, age 12:					
First stage				**0.294**	**0.329**
				0.067	0.066
# Police interactions age 12 (0/1)	0.041	0.007	0.004	-0.097	**-0.171**
		0.009	0.009	0.101	0.083
Police interactions age 12 (0/1)	0.033	0.003	0.001	-0.107	*-0.122*
		0.006	0.006	0.072	0.065
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		4556			4287
Crime outcomes, age 13:					
First stage				**0.294**	**0.329**
				0.067	0.066
# Police interactions age 13 (0/1)	0.080	0.025	0.029	-0.307	*-0.322*
		0.026	0.029	0.193	0.172
Police interactions age 13 (0/1)	0.045	0.001	0.002	-0.118	**-0.144**
		0.008	0.008	0.078	0.072
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		4556			4287

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Standard errors reported below coefficient estimates. Instrument lies in 0-0.6 interval. Standard errors clustered at cohortXcounty level. Control variables as in Table 2. One child excluded because of treatment before age 5.

VI.E *Heterogeneity in Results*

We finally investigate whether results vary across subgroups. We have performed analyses for boys, children born to mothers with more than a high school degree, children without a mental retardation diagnosis, children with a birthweight above 3,000 grams, children born to mothers who did not smoke during pregnancy, and children who are not in treatment with the five most common other types of nervous system drugs. Our main conclusions are robust to these exercises and the full set of results is available on request.

TABLE 4

DIFFERENCE-IN-DIFFERENCE RESULTS, EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE AGE 10 ON HEALTH OUTCOMES

	OLS, no control vars.	OLS control vars.	Age 4 comparisons
	(1)	(2)	(3)
	Health outcomes, age 10:		
# Hospital contacts, age 10	*-0.070*	**-0.074**	**-0.140**
	0.034	0.033	0.050
Hospital contacts, age 10 (0/1)	*-0.026*	*-0.027*	**-0.062**
	0.014	0.014	0.020
# Emergency visits, age 10	*-0.034*	*-0.030*	**-0.070**
	0.018	0.018	0.035
Emergency visits, age 10 (0/1)	-0.016	-0.014	**-0.047**
	0.013	0.013	0.018
Sample size		4556	

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Bold (italic) indicates significance at the 5% (10%) level. Standard errors assume homoscedasticity unless otherwise noted. One child excluded because of treatment before age 5.

As pointed out above, many more children are diagnosed early and treated pharmacologically in the later cohorts. Two competing hypotheses may explain this development. First of all, it is possible that diagnostic tools have improved in recent years and that diagnosed cases born in later cohorts suffer from ADHD to the same extent as children born in earlier cohorts. In this case we will expect the effects of pharmacological treatment to be the same in early and late cohorts. Secondly, it is possible that the group of treated has been broadened to include children with less severe symptoms. In this case we will expect the effects of pharmacological treatment to decline in late

cohorts. Table 5 investigates this for the larger health sample. We distinguish between the 1990-1994 cohorts and the 1995-1999 cohorts from the original sample and add interactions between the treatment indicator (and the instrument in the IV analysis) and a recent-cohort-indicator. We see that estimated effects are smaller in the more recent cohorts and significantly so in the panel data analysis. There are still significant gains from treatment in the later cohorts,[18] but the results are at least in line with diminishing returns to broadening the group of treated.

TABLE 5

EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE

AGE 10 ON HEALTH OUTCOMES IN EARLY AND LATER BIRTH COHORTS

	OLS	Difference-in difference		IV
	(1)	(2)		(3)
			Treatment	Treatment* (birth cohort > 1994)
			First stage	
Propensity to treat			**0.195**	-0.031
			0.085	0.019
Propensity to treat * (birth cohort > 1994)			0.165	**0.402**
			0.121	0.089
			Second stage	
# Hospital contacts, age 10				
Treatment	**-0.222**	**-0.271**		*-2.454*
	0.057	0.081		1.489
Treatment * (birth cohort > 1994)	**0.197**	*0.168*		2.088
	0.069	0.089		1.504
Hospital contacts, age 10 (0/1)				
Treatment	**-0.075**	**-0.122**		-0.429
	0.025	0.036		0.387
Treatment * (birth cohort > 1994)	**0.065**	**0.076**		0.325
	0.030	0.039		0.395
Control variables	YES	YES		YES
Sample size	4556	4556		4287

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Age 4 comparisons in column 3. Bold (italic) indicates significance at the 5% (10%) level. Standard errors clustered at cohortXcounty. Control variables as in Table 2. One child excluded because of treatment before age 5.

---

[18] If we run the analyses separately for the 1990-1994 and 1995-1999 birth cohorts, the panel data analysis estimates an effect of -0.122 (standard error 0.040) for the early birth cohorts and an effect of -0.046 (standard error 0.024) for the late birth cohorts.

VII. Conclusion

This paper investigates the effect of early pharmacological treatment of ADHD on children's human capital development while relying on register-based data on diagnoses and treatment for children born in 1990-1999. Our main analysis exploits variation in hospital level differences in propensities to prescribe treatment and a particular innovation of our paper is that we are to some extent able to control the severity of ADHD for the purposes of constructing the instrumental variable. Our analysis of effects on health outcomes allows for a complementary difference-in-difference analysis where we compare outcomes of treated children prior to (age four) and after treatment (age 10 +) with untreated diagnosed children before and after.

We find that the behavior of specialist physicians varies considerably across hospitals and that the prescribing behavior does affect the probability that a given child is treated. Results show that children diagnosed with ADHD on the margin of receiving pharmacological treatment have *fewer* hospital contacts if treated and are *less* likely to be charged with crime. We do find that estimated health effects are significantly smaller in later cohorts where more children are diagnosed and treated pharmacologically before the age of ten. There are still significant gains from treatment in the later cohorts, but these results are consistent with a hypothesis of diminishing returns to expanding treatment. This is particularly important in terms of policy-making because the share of children diagnosed in Denmark is relatively low and those diagnosed more likely to be higher in the severity spectrum than in other countries. Hence our study does not license much broader use of stimulants. Rather, it suggests that prescribing pharmacological treatment to children with less severe ADHD may not be beneficial in terms of the outcomes considered in this paper. In general, the degree to which broadening treatment is beneficial is likely to depend upon the prevailing treatment culture; in a setting where highly symptomatic children are deprived of treatment, granting access to treatment is likely to be more beneficial than in settings where treatments are given to children with mild or no symptoms.

From the point of view of the individual, the family and the society, the long-term benefits of pharmacological treatment of ADHD extend beyond the relief of the symptoms related to the syndrome. These benefits should be traded off against the (low) financial costs of the drugs combined with the potential detrimental short- and long-term side effects of medication such as insomnia, decreased appetite and increased blood pressure.

Literature

Blundell and Costa Dias (2009) 'Alternative Approaches to Evaluation in Empirical Microeconomics.' *Journal of Human Resources* 44, 565-640.

Coleman, J., E. Katz and H. Menzel (1957), 'The Diffusion of an Innovation among Physicians.' *Sociometry* 20 (4): 253-270.

Cunha, F., J. Heckman, L. Lochner and D. Masterov (2006), 'Interpreting the Evidence on Life Cycle Skill Formation', in E. A. Hanushek and F. Welch (eds.), *The Handbook of the Economics of Education*, New York, North Holland.

Cunha, F. and J. Heckman (2007), 'The Technology of Skill Formation', *The American Economic Review (Papers and Proceedings)* 97, 31-47.

Currie, J. (2011), 'Inequality at Birth. Some Causes and Consequences', *The American Economic Review (Papers and Proceedings)* 101, 1-22. DOI: 10.1257/aer.101.3.1

Currie, J. and D. Almond (2010), 'Human Capital Development before Age Five', in O. Ashenfelter and D. Card (eds.), *The Handbook of Labor Economics* Vol. 4b, Berkeley, Elsevier. DOI: 10.1016/S0169-7218(11)02413-0

Currie, J. and M. Stabile (2006), 'Child Mental Health and Human Capital Accumulation: The Case of ADHD, *Journal of Health Economics* 25, 1094-1118.

Currie, J., M. Stabile and L. E. Jones (2013), 'Do Stimulant Medications Improve Educational and Behavioral Outcomes for Children with ADHD? NBER working paper 19105.

Dalsgaard, S., M. K. Humlum, H. S. Nielsen and M. Simonsen (2012), "Relative Standards in ADHD Diagnoses: The Role of Specialist Behavior." *Economics Letters* 117, 663-665. DOI: 10.1016/j.econlet.2012.08.008

Dalsgaard, S., M. K. Humlum, H. S. Nielsen and M. Simonsen (2013), "Common Danish standards in prescribing medication for children and adolescents with ADHD." *European Child & Adolescent Psychiatry*. DOI: 10.1007/s00787-013-0508-5

Dalsgaard, S., H. S. Nielsen, J. Leckman and M. Simonsen (forthcoming), "Gender and Injuries Predict Stimulant Medication." *Journal of Child and Adolescent Psychopharmacology*.

Dalsgaard S., P.B. Mortensen, M. Frydenberg and P. H. Thomsen (2002), 'Conduct problems, gender and adult psychiatric outcome of children with attention-deficit hyperactivity disorder.' *British Journal of Psychiatry* 181: 416-21.

Ding, W., S. Lehrer, N. Rosenquist and J. Audrain-McGovern (2009), 'The Impact of Poor Health on Academic Performance: New Evidence Using Genetic Markers', *Journal of Health Economics* 28, 578-597.

Doyle, J. (2007), 'Child Protection and Child Outcomes: Measuring the Effects of Foster Care', *American Economic Review* 97, 1583-1610.

Doyle, J. (2008), 'Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care', *Journal of Political Economy* 116, 746-770.

Duggan, M. (2005), 'Do New Prescription Drugs Pay for Themselves? The Case of Second-Generation Antipsychotics.' *Journal of Health Economics* 24(1): -31.

Elder, T. E. (2010), 'The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates.' *Journal of Health Economics* 29, 641–656. DOI: 10.1016/j.jhealeco.2010.06.003.

Elder, T.E. and D.H. Lubotsky (2009), 'Kindergarten Entrance Age and Children's Achievement – Impact of State Policies, Family Background, and Peers.' *Journal of Human Resources* 44(3): 641-683.

Evans, W.N., M. S. Morrill and S. T. Parente (2010), 'Measuring Excess Medical Diagnosis and Treatment in Survey Data: the Case of ADHD among School-age Children.' *Journal of Health Economics* 29, 657–673. DOI: 10.1016/j.jhealeco.2010.07.005.

Faraone, S.V. and A.E. Doyle (2001), 'The Nature and Heritability of Attention Deficit/Hyperactivity Disorder', *Child and Adolescent Psychiatric Clinics of North America* 10, 299-309.

Faraone S.V., J. Sergeant, C. Gillberg and J. Biederman (2003), 'The Worldwide Prevalence of ADHD: is it an American Condition? *World Psychiatry* 2, 104-13.

Fletcher, J. and B. Wolfe (2008), 'Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited', *Journal of Health Economics* 27, 794-800.

Frazier, T.W., H.A. Demaree and E.A. Youngstrom (2004), 'Meta-analysis of Intellectual and Neuropsychological Test Performance in Attention-Deficit/Hyperactivity Disorder', *Neuropsychology* 18, 543-555.

Heckman, J. J. (2008), 'Schools, Skills, and Synapses', *Economic Inquiry* 46(3): 289-324.

Jensen, P.S. , S.P. Hinshaw and H.C. Kraemer et al. (2001), 'ADHD Comorbidity Findings from the MTA Study: Comparing Comorbid Subgroups' *Journal of the American Academy of Child and Adolescent Psychiatry* 40, 147-158.

Kvist, A. P., H. S. Nielsen and M. Simonsen (2013), "The Importance of Children's ADHD for Parents' Relationship Stability and Labor Supply." *Social Science & Medicine* 88, 30-38. DOI: 10.1016/j.socscimed.2013.04.001

Lechner, M. (2011), "The Estimation of Causal Effects by Difference-in-Difference Methods." *Foundations and Trends in Econometrics* 4(3), 165-224. DOI: 10.1561/0800000014

Lichtenstein, P., L. Halldner, J. Zetterqvist, A. Sjölander, E. Serlachius, S. Fazel, N. Långström, and H. Larsson (2012), "Attention Deficit Hyperactivity Disorder Medication and Criminality." *New England Journal of Medicine* 367:2006-2014. DOI: 10.1056/NEJMoa1203241

Linnet K.M., S. Dalsgaard and C. Obel et al. (2003) 'Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: review of the current evidence.' *J. Psychiatry.* 160(6):1028-1040.

Linnet, K.M., K. Wisborg, N.J. Secher, P.H. Thomsen, C. Obel, S. Dalsgaard and T.B. Henriksen (2009), *"Coffee consumption during pregnancy and the risk of hyperkinetic disorder and ADHD: a prospective cohort study." Acta Paediatr.*: 98(1):173-179.

Maestas, N., K. J. Mullen and A. Strand (2013), "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt", *American Economic Review* 103(5): 1797-1829. DOI: 10.1257/aer.103.5.1797

Mannuzza, S. and R. Klein (2000), 'Long-term prognosis in attention-deficit/hyperactivity disorder.' *Child and Adolescent Psychiatric Clinics of North America* 9, 711-726.

Marcus, S.C., G.J. Wan, H.F. Zhang and M. Olfson (2008), 'Injury Among Stimulant-Treated Youth with ADHD.' *Journal of Attention Disorders* 12(1): 64-69.

Mendez, L., P. Singh, G. Harrison, Y.-S. Huang, X. Jin and S.C. Cho (2011), "Academic outcomes in Asian children aged 8-11 years with attention-deficit/hyperactivity disorder treated with atomoxetine hydrochloride." *International Journal of Psychiatry in Clinical Practice* ; 15: 145–156.

Molina, B.S., S.P. Hinshaw, J.M. Swanson et al. (2009), 'The MTA at 8 years: prospective follow-up of children treated for combined-type ADHD in a multisite study.' *Journal of the American Academy of Child and Adolescence Psychiatry* 48, 484-500.

MTA Cooperative Group (1999), 'A 14-Month Randomized Clinical Trial of Treatment Strategies for Attention-Deficit Hyperactivity Disorder', *Arch Gen Psychiatry* 56: 1073-1086.

MTA Cooperative Group (2004), '24-Month Outcomes of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder.' *Pediatrics* 113: 754-763.

Munk-Jørgensen, P. and P.B. Mortensen (1997), 'The Danish Psychiatric Central Register', *Danish Medical Bulletin* 44, 82-84.

Obel, C., J. Olsen and T. B. Henriksen et al. (2011), 'Is maternal smoking during pregnancy a risk factor for Hyperkinetic disorder? – findings from a sibling design.' *International Journal of Epidemiology* 40: 338-345. DOI: 10.1093/ije/dyq185

Paykina, N. and L. Greenhill (2008), 'Attention deficit hyperactivity disorder.' In Findling R., ed., *Clinical manual of child and adolescent psychopharmacology* Washington, DC: American Psychiatric Publishing.

Solanto, M.V. (1998), 'Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: a review and integration.' *Behav. Brain Res* 94(1): 127-152.

Soumerai, S. B., T. J. McLaughlin, J. H. Gurwitz, E. Guadagnoli, P. J. Hauptman, C. Borbas, N. Morris, B. McLaughlin, X. Gao, D. J Wilson, R. Asinger and F. Gobel (1998), 'Effect of local

medical opinion leaders on quality of care for acute myocardial infarction: a randomized controlled trial.' *Journal of the American Medical Association* 6 (17): 1358-1363.

Sroufe, L. A. (2012), 'Ritalin Gone Wrong' *New York Times*. January 29, 2012.

The Danish Association for Child and Adolescent Psychiatry (2008), 'Reference Program for Diagnosing and Treating Children and Adolescents with ADHD' (Referenceprogram for udredning og behandling af børn og unge med ADHD).

Van der Oord S., P.J. Prins, J. Oosterlaan and P.M. Emmelkamp (2008), 'Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: a meta analysis.' *Clin. Psychol. Rev.* 28(5): 783-800.

WHO (1993),'The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research', World Health Organization.

Appendix

TABLE A1

OBSERVABLE CHARACTERISTICS AT CHILDBIRTH, CHILDREN WITH

AND WITHOUT AN ADHD DIAGNOSIS BEFORE THE AGE OF TEN

Variable	No ADHD diagnosis			ADHD diagnosis before age 10		
	# obs	Mean	Std. Dev.	# obs	Mean	Std. Dev.
Child:						
Boy (0/1)	710120	0.51	0.50	4557	**0.84**	0.37
5-minute APGAR score	654512	9.83	0.86	4331	**9.73**	1.10
Birthweight less than 1,500 gram	656115	0.01	0.08	4353	**0.02**	0.14
Birthweight, 1,500-2,500 grams	656115	0.04	0.20	4353	**0.07**	0.25
Birthweight, above 2,500 grams	656115	0.95	0.21	4353	**0.91**	0.28
Complications at birth (0/1)	710120	0.21	0.40	4557	**0.29**	0.45
Gestation length (weeks)	654637	39.59	1.92	4346	**39.21**	2.53
Mental retardation (0/1)	710120	0.00	0.06	4557	**0.11**	0.32
Mother:						
Age at childbirth	707357	29.05	4.81	4553	**28.14**	5.15
High school or less (0/1)	648886	0.39	0.49	4358	**0.52**	0.50
Length of education (years)	648886	12.25	2.52	4358	**11.41**	2.30
Unemployed less than 13 weeks	664613	0.83	0.38	4431	**0.80**	0.40
Unemployed 13-26 weeks (0/1)	664613	0.11	0.32	4431	**0.14**	0.34
Unemployed more than 26 week	664613	0.06	0.23	4431	**0.06**	0.24
Employed in November (0/1)	664576	0.63	0.48	4431	**0.51**	0.50
Gross income (2004 prices)	664576	188327	109197	4431	**174186**	74643
Psychiatric diagnosis (0/1)	707736	0.04	0.19	4553	**0.10**	0.30
Heart disease (0/1)	707736	0.03	0.16	4553	**0.03**	0.17
Respiratory disease (0/1)	707736	0.11	0.31	4553	**0.19**	0.39
Smoker (0/1)	562081	0.08	0.27	3945	**0.18**	0.39
Father:						
Age at child birth	684806	31.89	5.76	4277	**31.23**	6.25
High school or less (0/1)	632113	0.32	0.47	4064	**0.44**	0.50
Length of education (years)	632113	12.21	2.58	4064	**11.33**	2.37
Unemployed less than 13 weeks	651266	0.88	0.32	4197	**0.84**	0.36
Unemployed 13-26 weeks (0/1)	651266	0.06	0.23	4197	**0.08**	0.27
Unemployed more than 26 week	651266	0.06	0.24	4197	**0.08**	0.27
Employed in November (0/1)	651237	0.89	0.32	4197	**0.84**	0.36
Gross income (2004 prices)	651237	296103	199811	4197	**259101**	144851
Psychiatric diagnosis (0/1)	685964	0.03	0.18	4281	**0.08**	0.27
Heart disease (0/1)	685964	0.03	0.17	4281	**0.04**	0.20
Respiratory disease (0/1)	685964	0.09	0.29	4281	**0.14**	0.34

Notes: Bold indicates that mean for children with an ADHD diagnosis before the age of ten is significantly different from the mean for children without an ADHD diagnosis at the 5% level. With the exception of mental retardation, which is diagnosed in connection with the ADHD diagnosis, all variables are measured in the year just prior to the birth of the child or in connection with childbirth. Unemployed less than 13 weeks includes no unemployment.

TABLE A2

EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE AGE 10

		OLS		IV
	Mean	(2)		(4)
Crime outcomes, age 14:				
First stage			**0.294**	**0.297**
			0.067	0.071
# Police interactions age 14 (0/1)	0.156	0.021	0.023	-0.037	-0.164
		0.037	0.036	0.386	0.372
Police interactions age 14 (0/1)	0.077	-0.008	-0.008	-0.038	-0.048
		0.009	0.009	0.107	0.106
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		3719			3554
Crime outcomes, age 15:				
First stage			**0.258**	**0.249**
			0.077	0.077
# Police interactions age 15 (0/1)	0.201	**0.093**	**0.093**	-0.569	-0.600
		0.045	0.042	0.506	0.504
Police interactions age 15 (0/1)	0.101	*0.023*	*0.020*	-0.109	-0.127
		0.012	0.012	0.176	0.179
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		3055			2914
Crime outcome, age 16:				
First stage			**0.237**	**0.214**
			0.100	0.101
Charged age 16 (0/1)	0.040	0.000	0.002	-0.179	-0.158
		0.001	0.010	0.163	0.175
Cohort and regional dummies		YES	YES	YES	YES
Individual level control variables		NO	YES	NO	YES
Sample size		1938			1826

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Standard errors reported below coefficient estimates. Instrument lies in 0-0.6 interval. Standard errors clustered at cohortXcounty level. Control variables as in Table 2. Grades and teacher evaluations standardized at cohort level. One child excluded because of treatment before age 5.

TABLE A3

EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE AGE 10

SELECTED OUTCOMES, VARYING INSTRUMENTAL VARIABLE DEFINITIONS

		OLS	IV	IV less severe ADHD	IV county at diagnosis	IV county at birth
	Mean	(1)	(2)		(3)	(4)
			Health outcomes, age 10:			
First stage			**0.329**	**0.358**	**0.332**	**0.453**
			0.066	0.060	0.082	0.076
# Hospital contacts, age 10	0.498	*-0.027*	**-0.800**	*-0.601*	**-1.354**	**-0.704**
		0.014	0.441	0.354	0.641	0.436
Hospital contacts, age 10 (0/1)	0.278	**-0.074**	-0.166	-0.194	-0.278	-0.024
		0.033	0.156	0.122	0.200	0.144
Control variables		YES	YES	YES	YES	YES
Sample size	4556		4287	4256	4464	4522

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Instruments lie in 0-0.6 interval. Standard errors clustered at cohortXcounty level. Control variables as in Table 2. One child excluded because of treatment before age 5.

TABLE A4

DIFFERENCE-IN-DIFFERENCE RESULTS,

HOSPITALS WITH LOW AND HIGH TREATMENT INTENSITY

	Wards with low treatment intensity (< median)	Wards with high treatment intensity (> median)
	Health outcomes, age 10:	
# Hospital contacts, age 10	*-0.149*	*-0.114*
	0.086	0.060
Hospital contacts, age 10 (0/1)	-0.052	**-0.066**
	0.032	0.028
# Emergency visits, age 10	-0.038	**-0.069**
	0.044	0.035
Emergency visits, age 10 (0/1)	-0.023	**-0.052**
	0.028	0.024
Sample size	2131	2156

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Age 4 comparisons. Bold (italic) indicates significance at the 5% (10%) level. Standard errors assume homoscedasticity. One child excluded because of treatment before age 5. Treatment intensity (instrument) lies in 0-0.6 interval.

TABLE W1

PERCENTAGE OF BIRTH COHORT DIAGNOSED WITH ADHD BEFORE AGE TEN

Birth Cohort	Cohort size	Percent with ADHD among all	Average age at diagnosis	Percent with ADHD among boys	Percent with ADHD among girls
1990	69,026	0.29	7.01	0.48	0.08
1991	69,667	0.37	6.85	0.63	0.10
1992	72,869	0.42	6.93	0.70	0.12
1993	72,227	0.44	7.03	0.73	0.14
1994	74,766	0.53	7.11	0.88	0.16
1995	74,342	0.63	7.24	1.04	0.19
1996	71,700	0.72	7.20	1.17	0.25
1997	71,342	0.84	7.22	1.42	0.23
1998	69,549	0.96	7.22	1.52	0.36
1999	69,189	1.21	7.17	1.90	0.49
All cohorts	714,677	0.64	7.14	1.04	0.21

TABLE W2

PERCENTAGE OF CHILDREN DIAGNOSED WITH ADHD IN

PHARMACOLOGICAL TREATMENT FOR ADHD BEFORE THE AGE OF TEN

Birth Cohort	No. diagnosed	Percentage treated	Average age first treatment (in treatment before age 10)	Percentage treated among boys	Percentage treated among girls
1990	197	18.78	8.16	19.88	11.54
1991	257	19.84	8.14	21.43	9.09
1992	304	21.71	8.12	22.52	16.67
1993	318	22.33	8.14	22.30	22.45
1994	394	24.87	7.97	25.82	19.30
1995	468	28.21	8.19	30.08	17.39
1996	518	33.20	8.26	33.56	31.40
1997	599	34.39	7.93	35.07	30.00
1998	665	38.95	8.03	38.79	39.67
1999	837	43.61	7.93	45.54	35.76
All cohorts	4557	31.97	8.04	32.70	28.16

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten.

TABLE W3

OLS, CHILD AND PARENTAL BACKGROUND CHARACTERISTICS

AND HOSPITALS' PROPENSITY TO PRESCRIBE

	IV: Propensity to treat at ward	
	Coefficient Estimate	Standard Error
Child variables:		
Injury and poisoning* (DS00-DT98), age 0 (0/1)	0.004	0.016
Injury and poisoning* (DS00-DT98), age 1 (0/1)	-0.006	0.007
Injury and poisoning* (DS00-DT98), age 2 (0/1)	0.008	0.006
Injury and poisoning* (DS00-DT98), age 3 (0/1)	-0.004	0.006
Injury and poisoning* (DS00-DT98), age 4 (0/1)	-0.001	0.006
Boy (0/1)	-0.003	0.007
5-minute APGAR score	-0.020	0.023
Birth weight less than 1,500 grams (0/1)	*0.018*	0.025
Birth weight, 1,500-2,500 grams (0/1)	-0.016	0.011
Complications at birth (0/1)	0.007	0.007
Gestation length (weeks)	0.000	0.001
Mental retardation diagnosis (0/1)	0.003	0.009
Mother variables:		
Age at child birth	0.000	0.001
Length of education (years)	-0.002	0.001
Unemployed less than 13 weeks (0/1)	-0.003	0.007
Employed in November (0/1)	0.000	0.007
Gross income, kr. 100.000 (2004 prices)	**0.013**	0.005
Psychiatric diagnosis (0/1)	0.007	0.009
Heart disease (0/1)	-0.011	0.013
Respiratory disease (0/1)	-0.004	0.007
Smoker (0/1)	0.008	0.008
Father variables:		
Age at child birth	0.000	0.001
Length of education (years)	0.001	0.001
Unemployed less than 13 weeks (0/1)	-0.003	0.008
Employed in November (0/1)	-0.014	0.009
Gross income, kr. 100.000 (2004 prices)	0.001	0.002
Psychiatric diagnosis (0/1)	0.012	0.009
Heart disease (0/1)	0.020	0.013
Respiratory disease (0/1)	0.009	0.008
Sample size	4556	
R-squared	0.222	

Notes: Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Cohort and regional dummies included. Standard errors clustered at cohortXhospital level. One child excluded because of treatment before age 5.

TABLE W4

PROBIT, PROPENSITY TO DIAGNOSE AND PROPENSITY

TO PRESCRIBE (COUNTY LEVEL)

Variable	Marginal effect	Standard error
Propensity to prescribe	0.0014	0.0011
Controls	YES	
Sample size	714677	

Notes: Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Conditioning set corresponds to the variables in Table 2 . Cohort and regional dummies included. Standard errors clustered at cohortXcounty level.

TABLE W5

OLS, PROPENSITY TO PRESCRIBE AND THE FIVE MOST

COMMON OTHER TYPES OF NERVOUS SYSTEM DRUGS

	IV: Propensity to treat at ward	
	Coefficient Estimate	Standard Error
Other types of nervous system drugs:		
Valproic acid (N03AG01), antiepileptic/migrain	**-0.039**	0.016
Lamotrigine (N03AX09), anti-epileptic	0.015	0.017
Risperidone (N05AX08), antipsychotic	-0.006	0.011
Clopenthixol (N05AF02), antipsychotic	0.030	0.019
Diazepam (N05BA01), anxiety/insomnia	-0.005	0.011
Sample size	4556	

Notes: Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Conditioning set corresponds to the variables in Table 2. Cohort and regional dummies included. Standard errors clustered at cohortXcounty level. One child excluded because of treatment before age 5.

TABLE W6

SENSITIVITY ANALYSIS, MINIMUM HOSPITAL SIZE AND USE OF TREATMENT

	Ward > 1 child	Ward > 5 children	Ward > 10 children	Ward treats > 1 other child
	Health outcomes, age 10:			
First stage	**0.329**	**0.330**	**0.341**	**0.320**
	0.066	0.078	0.088	0.094
# Hospital contacts, age 10	**-0.800**	**-0.940**	*-1.040*	**-1.260**
	0.441	0.470	0.581	0.619
Hospital contacts, age 10 (0/1)	-0.166	-0.247	-0.231	-0.306
	0.156	0.188	0.213	0.214
Control variables	YES	YES	YES	YES
Sample size	4287	3905	3330	3870

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Bold indicates that coefficient is significant at the 5 % level and italic indicates significance at the 10 % level. Conditioning set corresponds to the variables in Table 2. Standard errors reported below coefficient estimates. Instrument lies in 0-0.6 interval. Standard errors clustered at cohortXcounty level. One child excluded because of treatment before age 5.

TABLE W7

EFFECTS OF PHARMACOLOGICAL TREATMENT BEFORE AGE 10 ON

HEALTH OUTCOMES, BY OBSERVABLE CHARACTERISTICS

	Boys	Mentally retarded excluded	Mothers more than high school	Birthweight > 3000 g	Non-smoking mothers	Not treated with other types of nervous system drug
IV analysis:						
First stage	**0.314**	**0.297**	**0.349**	**0.261**	**0.277**	**0.234**
	0.074	0.073	0.084	0.075	0.073	0.071
# Hospital contacts, age 10	**-1.286**	**-1.024**	-0.169	**-1.167**	*-1.154*	**-1.405**
	0.568	0.537	0.483	0.588	0.630	0.741
Hospital contacts, age 10 (0/1)	-0.258	-0.231	-0.162	-0.275	-0.152	-0.258
	0.187	0.190	0.209	0.223	0.205	0.241
Control variables	YES	YES	YES	YES	YES	YES
Sample size	3606	3804	2173	3346	3620	3751
Panel data analysis:						
# Hospital contacts, age 10	**-0.109**	**-0.122**	**-0.159**	**-0.125**	**-0.149**	**-0.128**
	0.044	0.057	0.055	0.051	0.046	0.052
Hospital contacts, age 10 (0/1)	**-0.052**	**-0.060**	-0.050	**-0.048**	**-0.066**	**-0.067**
	0.021	0.023	0.032	0.020	0.025	0.026
Control variables	YES	YES	YES	YES	YES	YES
Sample size	3829	4039	2300	3560	3831	3998

Notes: Treatment is defined as pharmacological treatment for at least six months in a year before the age of ten. Age 4 comparisons in panel data analysis. Bold (italic) indicates significance at the 5% (10%) level. Conditioning set corresponds to the variables in Table 2. Standard errors clustered at cohortXcounty. One child excluded because of treatment before age 5.