The Rockwool Foundation Research Unit

Study Paper No. 83

The Effects of Medical Treatment of Attention Deficit Hyperactivity Disorder (ADHD) on Children's Academic Achievement

Maria Keilow, Anders Holm and
Peter Fallesen

Copenhagen 2015

The Effects of Medical Treatment of Attention Deficit Hyperactivity Disorder (ADHD) on Children's Academic Achievement

---

Study Paper No. 83

Published by:

Address:
The Rockwool Foundation Research Unit
Soelvgade 10, 2.tv.
DK-1307 Copenhagen K

Telephone	+45 33 34 48 00
E-mail	forskningsenheden@rff.dk
web site:	www.en.rff.dk

ISBN 978-87-93119-24-6
9788793119246
ISSN 0908-3979

March 2015

The Effects of Medical Treatment of Attention Deficit Hyperactivity Disorder (ADHD) on Children's Academic Achievement

Maria Keilow
SFI — The Danish National Centre for Social Research

Anders Holm*
Department of Sociology, University of Copenhagen
SFI — The Danish National Centre for Social Research
CAM – Centre for Applied Microeconometrics

Peter Fallesen
Rockwool Foundation Research Unit

ABSTRACT

We use Danish register data to estimate the effect of medical treatment of ADHD on children's academic achievement. Using a sample of 7,523 children who undergo medical treatment, we exploit plausibly exogenous variation in medical nonresponse to estimate the effect of medical treatment on school-leaving grades. Heckman two-stage sample selection models allow us to account for selection into the sample of children treated medically for ADHD. We find significant effects of treatment on ninth grade school-leaving grade point average (GPA). Compared to consistent treatment, part or full discontinuation of treatment has large significant negative effects on teacher evaluation and exam GPA, reducing grades with .18 to .19 standard deviations. A supplementary identification strategy and placebo regressions support our findings. The results demonstrate that ADHD treatment may mitigate the negative social consequences of behavioral problems.

*Corresponding author: ah@soc.ku.dk. Øster Farimagsgade 5A, Byg.16, 1353 København K, Denmark. The study received support from DSF #10-093105. The authors also thank the Rockwool Foundation for additional support.

INTRODUCTION

A large literature has established that physical health problems affect children's educational attainment negatively (e.g. Conti, Heckman and Urzua 2010; Jackson 2009, 2010). Yet, we know less about the link between mental health and education—untreated mental health problems not directly associated with cognitive abilities may still have a severe impact on children's learning potential. In particular, behavioral disorders that affect attention, such as Attention Deficit/Hyperactivity Disorder (ADHD), could have a detrimental impact on educational attainment if left untreated. On average, children diagnosed with ADHD have a lower grade point average (GPA), do worse on tests, have higher retention rates and absenteeism, and have lower high school and college completion rates (Biederman et al. 2004; Currie and Stabile 2004; Eisenberg, Golberstein and Hunt 2009; Fletcher and Wolfe 2007; Kessler et al. 1995; Mannuzza et al. 1993, 1997; McLeod and Kaiser 2004). If ADHD symptoms directly affect educational outcomes negatively, successful treatment could reduce such negative effects.

Although the use of social non-medical interventions and cognitive therapy to treat ADHD have increased, the predominant and most well documented treatment form that reduce core symptoms among school-aged children remains pharmacological treatment (Boyle and Jadad 1999; Danish Health and Medicines Authority 2014; MTA Coorporate Group 2004; Wilens and Spencer 2000). Yet only few studies on the subject of ADHD and education consider long-term individual effects of medical treatment on educational outcomes (exceptions are Currie and Stabile 2004; Currie, Stabile and Jones 2013; Dalsgaard, Nielsen and Simonsen 2014). These studies indicate that medication has no long-term effects on academic achievement in countries with lower treatment rates (Dalsgaard et al. 2014) and might even have negative effects in countries with high shares of children receiving treatment (Currie and Stabile 2004; Currie et al. 2013).

Previous research has focused on the effect of treatment on the marginal child—that is, what is the effect of expanding (or reducing) the share of children receiving treatment. Yet such local average treatment effects (Angrist and Imbens 1994) may be vastly different from the average effect of treating a child with ADHD. In addition, the previous work does not consider what extent the efficiency of treatment may moderate the impact on educational outcomes. Nor do they consider variation in treatment patterns subsequent to treatment initiation, i.e. if the individual treated receives full or only partial treatment.

Medical studies show that 25-30 percent of treated individuals with ADHD are nonresponders who show either no response to the medication or have severe side effects, both cases causing them to discontinue treatment (Barkley 1977; Schachar and Sugarman 2000; Spencer et al. 1996). Moreover, medical studies document only minor differences in background characteristics between responders and nonresponders (Barkley 1977; Powell et al. 2011; Spencer et al. 1996; Wilens and Spencer 2000). Comparing the academic achievement outcomes of responders and nonresponders thus allows us to estimate the average effect of medical ADHD treatment on academic achievement.

This study uses variation in medical response to beginning pharmacological treatment of ADHD to estimate causal effects of different pharmacological treatment patterns on the academic achievement of students treated for ADHD. In particular, we use purchases of ADHD medication to determine whether respondents discontinue treatment, or exhibit either ambiguous or consistent treatment patterns. Estimating inverse treatment effects, the empirical analysis thus tests if complete or partial discontinuation of treatment has a negative impact on academic achievement compared to continuous treatment.

Following this approach, our study expands previous work that compare medically treated children diagnosed with ADHD to non-treated children diagnosed with ADHD (e.g. Dalsgaard et al. 2014). These studies ignore whether treated individuals discontinue treatment or not. By investigating the effects of treatment within a sample of children who all begin pharmacological ADHD treatment, we circumvent problems related to initial selection into medical treatment. For the sample used in this study, all individuals have chosen to commence medical treatment.

We further test our main identification strategy by estimating a placebo regression of pharmacological treatment patterns on students from the same cohorts as in the main analysis, i.e. students who initiate ADHD treatment *after* they receive their school-leaving GPA. The analysis shows that our main identification strategy is robust to misspecifications and that medical treatment has a sizeable effect on the long-term academic achievement outcome of GPA. The effect is comparable to the gender differences in GPA.

The remainder of this paper is organized as follows: Section 2 reviews the literature on the effect of ADHD and ADHD treatment on educational outcomes. Section 3 presents the data for our empirical analysis. Section 4 presents our identification strategies and our statistical method. Section 5 presents results. Section 6 discusses the findings and concludes.

ADHD, MEDICAL TREATMENT, AND EDUCATIONAL OUTCOMES

Children diagnosed with ADHD on average attain 2.2 to 2.5 years less schooling than non-ADHD peers do, and 25 percent of students with ADHD drop out of high school (Currie and Stabile 2004; Fletcher and Wolfe 2007; Mannuzza et al. 1993, 1997; McLeod and Kaiser 2004; Miech et al. 1999). Research shows that exhibiting internalizing and externalizing behaviors associated with ADHD at age six to eight reduces the likelihood of obtaining a high school degree significantly and may reduce the likelihood of attending college for high school graduate students (McLeod and Kaiser 2004). Overall, behavioral disorders substantially impair children's educational attainment (McLeod, Uemura and Rohrman 2012; Miech et al. 1999).

Children with ADHD also display lower cognitive achievements. Using GPA and test scores as indicators, Currie and Stabile (2004) find that moderate to severe general ADHD symptoms reduce future reading and math scores. Correspondingly, Rodriguez et al. (2007) find significant association in Danish, Finnish and Swedish cohorts between core ADHD symptoms and scholastic impairment in reading, writing, and mathematics. Particularly, inattention was related to a two-fold to ten-fold increase in scholastic impairment. Distinguishing between attention deficits and externalizing problems, a recent study (Breslau et al. 2009) shows that whereas attention problems at a young age predict poorer math and reading achievement, the influence of externalizing and internalizing problem behaviors have little to no impact. Similarly, Duncan et al. (2007) find that early attentional deficits, more than externalizing behavior, appear to affect cognitive achievement.

Effect of medical treatment of ADHD on educational outcomes

Stimulants are the primary pharmacological therapy used for school-aged children diagnosed with ADHD and have been for several decades (Barkley 1977; Vaughan, March and Kratochvil 2012). Overall efficacy of stimulants and non-stimulants on core ADHD-symptoms is well documented for the age group 6-18 years by over 200 clinical RCT's completed since the 1960s (see e.g. Vaughan et al. 2012).[1]

Short-term medical treatment effects on academic achievement outcomes have been studied extensively and show positive effects (Carlson and Bunner 1993), whereas long-term studies are scarce and the effects are generally more diverse, smaller, or less significant than the short-term effects. Part of the reason for this may be methodological problems such as selection in and out of studies, lack of random assignment, small sample sizes, non-clinical settings, medical nonresponse issues unaccounted for and a wide variation in medication dosages (Carlson and Bunner 1993; Pelham and Smith 2000; Sprague and Sleator 1977).

Recent follow-up studies on a large-scale RCT, the *Multimodal Treatment Study of ADHD* (MTA Cooperative Group 1999), found no impact of medication on reading achievement three years after randomization, but also found that medication use had decreased by 62 percent after the study ended (Molina et al. 2009; MTA Coorporate Group 2004). Conversely, Hechtman et al. find significant improvement of various academic performance outcomes for stimulant responsive children with ADHD (without learning or conduct disorders) across treatment with medication alone and a treatment combination of medication and nonspecific psychosocial treatment (*attention control*) (Hecktman et al. 2004). These effects are persistent over at least two years.

Overall, despite significant short-term effects, most long-term studies show no medical effect on academic or cognitive outcomes, and those that do, conclude that effects are limited or that medical treatment may be most effective when combined with other intervention types (Carlson and Bunner 1993). However, they point to the methodological problems as the most probable explanation for the lack of empirical support for long-term effects and conclude that the want of research on this matter is substantial (Carlson and Bunner 1993:10). In sum, there seems to be only scant evidence of medical effects on long-term educational outcomes of children diagnosed with ADHD, which warrants studies that address this lack in the literature.

DATA

Data sources

We use data from Danish administrative registers collected by Statistics Denmark. The primary data sources are the Danish Registry of Medicinal Product Statistics (RMPS) and data from the national registers on educational attainment. Danish registers offer extensive individual-level information on all educational courses commenced in the public educational system and various academic performance measures. From 2002, the registers contain *teacher evaluation grades*[2] and *exam grades* of students who graduate the compulsory level of schooling (ninth grade in Denmark). From these grades, we create our outcome measure by calculating and standardizing exam and teacher evaluation GPA within the total population of children in Denmark.

Data from RMPS contain detailed information on legal purchases of prescription drugs in Denmark. The prescription data are linked to the individual identification number of the patients throughout the Danish health insurance system (see Kildemoes, Sørensen and Hallas 2011 for data description). The most apparent advantage of using these data is the large numbers of observations available, which is especially important when studying a subpopulation like that of children diagnosed with ADHD. Additionally, by using register data we avoid bias from refusal of participation, which is often a problem for studies of ADHD based on survey data (e.g. Boyle and Jadad 1999; MTA Cooperative Group 1999), and we avoid potential recall bias or reliability problems from self-reported information on medical treatment patterns.

Sample criteria

We sample children diagnosed with ADHD from the total population of 618,022 children who completed Danish compulsory schooling, i.e. ninth grade, from 2002 to 2011.[3] To identify children diagnosed with ADHD within this population, we initially select all children who redeemed at least one prescription for any type of ADHD medication (methylphenidate, atomoxetine, or modafinil) in the available RMPS data (years 1995-2011). We obtain a subsample of 12,486 children diagnosed with ADHD.[4] Of these children, some commence medical ADHD treatment prior to ninth grade exams (*treatment sample)*, whereas some start treatment subsequently (*placebo sample)*. We restrict both samples further using the following three criteria. First, children who either end medical treatment within the first three months of the RMPS data period or begin treatment during the last three months of the data period are excluded from analyses, as the treatment patterns of these children cannot be convincingly determined (n = 647). Second, the sample is restricted to cases with valid information on main outcome variables, i.e. teacher evaluation and exam GPA (n = 3146). Third, children who commence treatment around the time point of exams are excluded because the direction of causality is questionable, as they might commence treatment because they receive a low GPA (n = 1170). These restrictions result in a final sample of 7,523 children with a treatment sample consisting of 3,738 individuals from the birth cohorts 1984-1996, and, correspondingly, a placebo sample of 3,785 individuals from the cohorts 1983-1994.

Operationalization of treatment patterns

We determine individual medical treatment patterns using the longitudinal data from RMPS. In an actual RCT experiment, we would ideally divide the sample into just two groups, having individuals subjected to either treatment or non-treatment, corresponding to either discontinuation of treatment or consistent treatment. However, the variation of medical treatment patterns within the sample of individuals diagnosed with ADHD makes a binary distinction less clear-cut. Whereas some individuals end treatment after only a few purchases of medicine, others follow consistent treatment patterns for several years. Between these two outer extremes, data reveals a group of individuals with unstable treatment patterns interrupted by months without treatment. Hence, we define three types of pharmacological treatment patterns of use in the analysis: *Discontinued Pharmacological Treatment* (DPT), *Ambiguous Pharmacological Treatment* (APT), and *Continuous Pharmacological Treatment* (CPT).

The definition of discontinuation (DPT) follows strict restriction rules: we define DPT as having purchased medication for a maximum of three months within the data window, which just allows for initial medicine trial and dose titration. Likewise, the definition of continuous treatment (CPT) is restricted to those who have a regular and stable use of medication with purchases being no more than three months apart. The patterns of the remaining group, the APTs, are then per definition ambiguous as they cannot be convincingly placed in either the DPT or CPT category. All three categories are mutually exclusive. In all analyses, CPT is reference group, whereby we investigate inverse treatment effects of DPT and APT (discontinuation and ambiguous treatment) compared to CPT (a stable continuous treatment course).

We also use a rich set of control variables. Individual level information includes gender, and immigrant status, birth weight (in grams) and gestational length (in weeks). Studies show that birth characteristics predict ADHD symptoms (see Heinonen et al. 2010; Linnet et al. 2006) and correlate with academic achievement such as GPA (Fletcher and Lehrer 2009:15). To account for any effect of the duration of treatment, severity of symptoms, or potential correlations between age and treatment response (Pelham and Smith 2000:194), we control for individual age at the start of medical treatment. We also include a dummy for being graded with the new Danish grading scale introduced in 2008, as this change may affect the level of grades (Ministry of Children and Education 2007).

We measure parent characteristics one year prior to the birth of the respondent to avoid any post treatment confounding. Having a child with severe ADHD may potentially affect family resources such as income, unemployment, educational level attained, and family stability (Kvist, Nielsen and Simonsen 2013). Personal yearly income in USD (excluding any unearned or investment income) of the mother and father is included and indexed for 2009 prices. We include dummies for negative income for each parent as well, and variables identifying the proportion of the year that each parent was unemployed (on a scale from 1-1000). We specify parental educational level as ISCED code dummies (UNESCO Institute of Statistics 2012). We also include parental age at the birth of their child and dummy variables for whether either parent has an ADHD diagnosis. ADHD has genetic components, and studies estimate heritability up to 76 percent (Faraone et al. 2005; Vaughan et al. 2012). In addition, parents diagnosed with ADHD may execute parenting styles that more likely promote ADHD symptoms in their children. Table 1 presents the descriptive statistics of all background variables included in the analysis. We show data for the population of 2002-2011 graduates and for our restricted sample across treatment status.

* Table 1 about here *

Children in the treatment sample have lower gestational age (and may be more likely to be born preterm). The proportion of males in the treatment sample (74 pct.) is notably larger than in the population (50 pct.). The proportion of males is also higher in the placebo sample (58 pct.) than in the population data but also lower compared to the treatment sample. The proportion of non-immigrant Danes is larger in both the treatment sample and the placebo sample compared to the population data.[5] The overrepresentation is consistent with clinical accounts of diagnostic patterns. Looking at parental characteristics, parents of children in the treatment and the placebo sample are slightly younger, more likely to have an ADHD diagnosis, have higher unemployment rates, the mean paternal income is lower, and the educational variables indicate a greater proportion of parents with shorter length educations. These differences could indicate selection problems and thus incomparability between the placebo and the treatment sample. However, as we show in the results section, these differences do not seem to affect how medical treatment affects the outcome.

Table 2 presents descriptive statistics on outcome measures. We use GPAs from ninth grade exams and from teacher evaluations. Teacher evaluation grades reflect the individual academic performance (oral and/or written) of the student in class as evaluated by the teacher. The student's teacher as well as an external examiner assesses exam grades. Grades are assessed using the Danish 7 point grading scale, which corresponds to the ECTS-scale (Education and Culture DG 2009). However, all empirical models use GPAs that are standardized within the population sample.

* Table 2 about here *

For the population of compulsory school graduates from 2002-2011, 93 percent of all children have valid information on exam GPA and teacher evaluation GPA, compared to 73 percent of children diagnosed with and treated for ADHD. However, within the group of children initiating treatment there are only minor differences between treatment groups in how many students receive grades. Thus, our design circumvents this selection problem.[6]

METHOD

We estimate the effect of medical treatment on children diagnosed with ADHD on end of compulsory school GPA using a sample of children diagnosed with ADHD who initiate treatment. Using the difference between those continuing and those discontinuing medical treatment we recover the effects of medical treatment. Estimating credible causal effects in the absence of successfully randomized treatment is a daunting task and requires strong assumptions. Therefore, we elaborate below why our approach is a feasible way of identifying causal effects of medical treatment of ADHD and we discuss ways of testing key assumptions.

Our contribution has some advantages compared to prior studies claiming to recover causal effects of medication. Under the assumptions of our identification strategies, we estimate the average treatment effect on the treated. As mentioned in the introduction, this is, for example, opposed to (Currie et al. 2013), who use variation in time in the propensity to prescribe medication to credibly estimate causal effects. Currie et al. thus estimate the local average marginal treatment effect (see Angrist and Pischke 2009), i.e., the effect for individuals on the margin of being treated. Their study is silent, however, on the average treatment effect, which may vary considerably from the marginal treatment. Since the medical ADHD treatment regime has been increasingly inclusive (Connor 2002; Cuffe et al. 2001; Kessler et al., 1994, 2005; Polanczyk et al., 2007; Visser et al., 2010), this point must be addressed. As previously mentioned, other studies of ADHD treatment effects on educational outcomes have found effects, which can be interpreted as average treatment effects on the treated. However, these studies rely on weaker identification assumptions than found in our contribution, i.e. they use only simple controls for observed characteristics and no effort is made to further check for confounders nor check the robustness of the identification strategy.

Our main identification strategy exploits evidence from medical research showing that individual response to ADHD medication regarding the improvement of core symptoms is arbitrary with respect to the individual characteristics of the patient. In other words, *no* individual characteristics or baseline measures can predict a patients' response or nonresponse accurately and reliably in terms of the effect of medication (Barkley 1977; Pelham and Smith 2000; Powell et al. 2011; Spencer et al. 1996; Wilens and Spencer 2000). Consequently, among a group of individuals diagnosed with ADHD, whether or not the medication actually reduces symptoms of the diagnosed individual is, from a medical perspective, seemingly random. However, as we showed previously, treatment dropout is not merely a binary event – many initial prescribers to ADHD medication follow an ambiguous and inconsistent treatment pattern.[7] We therefore divide initial prescribers into three different groups; children with a consistent pattern of treatment (CPT), children with an ambiguous pattern of treatment (APT) and children with a discontinued pattern of treatment (DPT) in order to conduct our empirical analysis.

Analytical Setup

To measure the effect of medication on students' school-leaving GPA we run the following regression:

$$y_i = \beta x_i + \gamma_1 DPT_i + \gamma_2 APT_i + e_i \qquad (1)$$

where y_i is student GPA, x_i is a vector of individual characteristics, β is a vector of corresponding regression coefficients, DPT and APT are binary dummy indicators of DPT and APT patterns and e_i is the error term. Under random medical nonresponse, the regression coefficient γ_1 measures the average inverse treatment effect of the treated. This is because it measures the average difference in outcomes between those continuing medication (CPT) and those discontinuing medication (DPT) from the onset of diagnosis and until completing compulsory school, where GPA is measured.

The interpretation of γ_2 is less clear-cut, but captures the average difference in outcome between a continued consumption pattern and the average ambiguous treatment pattern (APT). The effect thus depends on both the effectiveness of the medication (generating a difference between those with an ambiguous and a continuous medical consumption) and on the distribution of the ambiguous treatment patterns in the data. That is, if medication is effective, the difference between APT and CPT and DPT depends on how "ambiguous" the APT's are. If they on average are close to full compliance, they should have, on average, the same outcome as the CPT's. On the other hand, if they, on average, are close to a fully disrupted treatment they should, on average, have the same outcome as the DPT's. At any rate, a negative sign of γ_2 may be taken as evidence that medication is effective.

Identifying a causal effect using differences in outcomes between DPT and CPT (and partially between APT and CPT) rests on the assumption that there is a one–to-one relationship between medical nonresponse and discontinuation of medical consumption, because we only observe discontinuation in the data. However, as mentioned in the introduction, even though random nonresponse is supported by empirical evidence, drop out from medical treatment in observed data may be nonrandom for a number of other reasons. Take up rates of prescriptions have a social gradient due to economic constraints for poorer families to finance medication consistently, and at-risk families may find it harder to follow medical advice or to comply with medical treatment despite negative side effects. Studies find that this applies for health care in general (see Adler and Newman 2002; Dunlop, Coyte and McIsaac 2000) and for mental health care and ADHD treatment specifically (e.g. Swanson 2003; Wang 2007). Such conditions may also help explain the existence of APT in the data. Some individuals may find it hard to follow medical treatment despite a favorable outcome like academic achievement.

However, for differences in background characteristics to be important with respect to estimating causal effects, background characteristics should both correlate with the allocation into DPT, APT and CPT and simultaneously have a direct effect on the outcome (see Alwin and Hauser 1975; Stolzenberg 1980).

To evaluate the extent that omitted variables bias our results we estimate eq. (1) both with and without observed covariates. Doing so enables us to see how much the inclusion of observed characteristics affect the estimates of the differences between CPT, DPT and APT. If the differences are unaffected by including observed covariates, this indicates either that background characteristics are uncorrelated with our treatment indicators or that they are unimportant for our outcome of interest. Either way, lack of change in the estimated differences indicates a causal interpretation of the estimated differences between CPT, DPT and APT, because it is then equally likely that unobserved characteristics also do not affect the differences in outcomes between CPT in each model.

Because we match data across cohorts, we also control for selection into our sample. During the last 20 years, the number of different medical products available on the Danish market has increased, while existing medical products have been enhanced (Danish Health and Medicines Authority 2010; Pottegård et al. 2012).[8] This development may increase the probability that individuals continue medical treatment for ADHD once they start. However, the observed changing rates of DPT, APT, and CPT across time may also reflect a change in the composition of students in our data.[9] To allow for this, we use a two-stage Heckman selection estimator to correct for sample selection into our data (Heckman 1977). Figure 1 shows the changing rates of DPT, APT, and CPT across our sample period.

* Figure 1 about here*

We use, as exclusion restrictions, children's residential region at age three and their birth cohort. Because the propensity to prescribe medication differs across national regions (Baumgardner et al. 2010; Bruckner et al. 2012; Fallesen and Wildeman Forthcoming), region of residence at a young age likely affects the propensity to enter medical treatment. As medical treatment has become increasingly more widespread, birth cohort may also affect the propensity to enter medical treatment. We control for region of residence at graduation in the main equation to allow for differences in grading practices and the effect of unobservables across regions. Birth cohort does not affect grades because grades are by construction independent across cohorts in accordance with the grading legislation of the Danish compulsory schools (Ministry of Children and Education 2013).[10]

Further, to test the exogeneity assumption of DPT (and APT) as an indicator of nonresponse to medical treatment, we use the availability of a placebo treatment group: students who commence medical treatment *after* graduation (thus after grades are measured) and for whom we should expect no treatment effect on their grades. Estimating eq. (1) on the placebo sample yields no effect unless nonresponders (DTP) are selected on unobservables compared to responders (CPT) and partial responders (APT). Obviously, the placebo and the treatment sample differs in the timing of their treatment, and this may imply underlying differences in their background characteristics.[11] However, as we show in the results section the differences across DPT, APT and CPT are unaffected by the inclusion of covariates in the treatment sample. Hence, it is unlikely that these differences are important when comparing DPT, APT and CPT across the treatment and placebo sample. Moreover, we show that students in the treatment sample and placebo sample are equally negatively affected by having an ADHD diagnosis. In summary, we provide two identification strategies for interpreting the difference between those disrupting and those continuing medical treatment as an average causal treatment effect of medical treatment on student GPA.

RESULTS

In this section, we present our results estimating eq. (1) on the sample of students diagnosed with ADHD. We present two sets of results, namely results for exam GPA and results for teacher evaluated GPA, both graded at the end of 9th grade in the end of Danish compulsory school. Table 3 shows estimation results for standardized exam GPA. We have replicated all treatment sample estimations correspondingly in the placebo sample. However, as these results are almost identical we only show the more statistically efficient estimates without adjustment for selection. In commenting on the results, we focus on our variables of interest and only briefly touch upon the estimated effects for the control variables.

* Table 3 about here *

Table 3 (column 2) reports significant differences across DPT, APT and CPT, such that children who follow continuous treatment (CPT) have significantly higher exam grades compared to those who discontinue treatment fully or partially (DPT and APT). The difference is more than one-tenth standard deviation between DPT and CPT and a tenth standard deviation between APT and CPT. Including control variables (column 3) further increases this difference, albeit not significantly. Controlling for background characteristics we thus find almost a fifth standard deviation difference between DPT and CPT and more than a tenth of a standard deviation difference between APT and CPT. Controlling for selection into our sample of initially treated students (column 4) makes virtually no differences in estimated differences across treatment groups.

For the placebo sample (column 5), we find no differences across treatment groups, indicating no selection bias on unobservables for the placebo sample. If we are willing to assume that this also applies for the treatment sample—supported by the fact that, for the treatment sample, there are no effects for including controls—this supports an interpretation of the differences between DPT, APT, and CPT as a causal effect of medication.

Inspecting the effect of the covariates on the outcome, we find many of the expected sizes and signs: Males and immigrants have lower exam GPA. Children whose mother is diagnosed with ADHD have a higher GPA, though no difference is seen if their father is diagnosed. Where this, at first glance, appears counterintuitive, it is likely due to selection of mothers being diagnosed—ADHD is hereditary, and mothers diagnosed are likely more able to address their children's needs because they themselves are undergoing or have undergone treatment. Parent educational level is positively correlated with exam GPA.[12] Comparing the estimates of the controls for the treatment sample and the placebo sample reveals similar sizes of the estimates for most control variables. This indicates that the selection process of the two samples of students diagnosed and initiating treatment for ADHD is similar. Had they been different, then we would have seen a marked differences between the estimates since different selection processes would have yielded different selection bias in the estimated coefficients.

Next we turn to the results for teacher evaluated GPA. This measure reflects both the academic achievement of the student as well as how desirable the student behaves in class. In this respect, teacher evaluated GPA reflects non-cognitive skills over and above the cognitive skills reflected in exam GPA. Table 4 presents results for standardized teacher evaluated GPA. Models are similar to those for exam GPA in table 3.

* Table 4 about here *

From table 4, we find comparable results for teacher evaluated GPA to those found from exam GPA. Treatment estimates are of similar magnitude compared to the estimates found for exam GPA. In addition, the control variables have similar effects on both GPA outcomes. Overall, this leads us to conclude that the effect of ADHD medication mainly works through student academic performance and not on behavioral traits over and above those needed to improve academic performance. However, we do find some small differences between the estimated influences of control variables on the outcomes. Mother's education has a somewhat stronger effect on teacher evaluated GPA than on exam GPA. This may reflect that the effect of "cultural capital" works more on teachers' assessment of the student than through pure academic achievement (Jæger 2011).

In sum, we find evidence of an effect of medical treatment on GPA for students diagnosed with ADHD who initiate medical treatment but have different subsequent treatment patterns. Students who disrupt treatment completely have almost one fifth of a standard deviation lower exam GPA and teacher evaluation GPA compared to students with continuous treatment pattern. Students with an ambiguous treatment pattern also obtain lower school-leaving GPA's than students with a continuous treatment pattern, although they still obtain a higher GPA than those who fully disrupt medical treatment. Our results are robust to the inclusion of control variables and selection into the group of treated students. Moreover, there are no differences in GPA outcomes across treatment patterns for children beginning treatment after receiving their school-leaving GPA, which supports our main results.

DISCUSSION

In this study, we have found that increased efficiency of medical treatment of ADHD can alleviate a substantial part of the GPA gap between children diagnosed with ADHD and their peers. Our empirical analyses show substantial and significant negative effects from discontinued pharmacological treatment (DPT) and from ambiguous pharmacological treatment (APT) on school-leaving teacher evaluation and exam GPAs compared to consistent treatment (CPT). Treatment effects are synonymously negative, in the sense that both DPT and APT lower grades across various model estimations and are robust to the statistical control for a large set of covariates and selection models.

Our findings support an overall beneficial effect of medical treatment on long-term individual academic achievement of children diagnosed with ADHD, a finding relevant for psychology, psychiatry, medicine, health economics, public health, and educational sociology. Children included in our study likely have moderate to severe symptoms, and for this sample, benefits may exceed any negative effects of treatment. In addition, Danish treatment rates of ADHD are more than 80 percent lower than, for example, US treatment rates (see Fallesen and Wildeman *forthcoming*). Although due consideration should be given to the necessity of medicine versus alternative modes of interventions or combined treatment, medication may help certain children diagnosed with ADHD to function in an everyday school context when baseline-treatment levels are of a moderate size.

As medical treatment affects ninth grade GPA of individuals diagnosed with ADHD, treatment likely influences subsequent educational and vocational tracks as well; for example choice of post-secondary tracks, dropout rates, unemployment propensity, adult profession, and adult income. International studies have found evidence of such effects from ADHD diagnosis. However, due to interdisciplinary gaps between social and medical studies in this field of study, research is needed that can adequately account for diagnosis *and* treatment as well as for the endogeneity and bias when studying long-term social outcomes of children diagnosed with ADHD.

This study found that medicine exerted a profound influence on the educational outcomes of children diagnosed with ADHD. Methodologically, the analysis is conditioned on the stable unit treatment value assumption, implying that potential individual outcomes are unrelated to the treatment status, outcomes or characteristics of other individuals. However, children with ADHD are part of social contexts—school classes, families, etc.—and the individual effects of treatment presented should ideally be construed in its context. Therefore, additional research should study the impact of ADHD treatment on peers and families. A few recent studies suggest that ADHD and medical treatment hereof affects academic outcomes of non-ADHD siblings (e.g., Fletcher and Wolfe 2007) and classroom peers (e.g., Aizer 2008). Moreover, current political focus on educational principles of inclusion of ADHD students in ordinary teaching environments calls for investigations of how these children and their peers perform academically, when included in regular classroom.

More than four decades of sociological research has voiced concerns about medicalization, social control, and the social construction of problem behaviors (e.g., Conrad 2006; Conrad and Schneider 1992). Recent research shows that increasing medical treatment rates of ADHD in populations with high baseline-treatment levels likely has adverse consequences for children (e.g., Currie et al. 2013). Yet, it remains a well-documented fact that children with behavioral and mental health problems, such as ADHD, have poorer later life outcomes than their peers. In addition, research from countries with low baseline-levels of ADHD treatment consistently shows that in this context, medical treatment lowers problem behavior (e.g., Dalsgaard et al. 2014). Our study adds to this research and our results indicate that medical treatment can improve the educational trajectories of children diagnosed with ADHD.

ENDNOTES

[1] New types and formulas of stimulants and non-stimulants have been introduced, resulting in improved management methods (see Madaan et al. 2006). RCT studies find that depot formulations result in better compliance rates among patients with ADHD (Sanchez et al. 2005). Pharmaceutical developments create a greater market supply of medical products available, which in turn increases the likelihood of finding suitable medical treatment for ADHD (Madaan et al. 2006; Pliszka 2003). Consequently, later born cohorts of ADHD diagnosed children have a higher probability of receiving efficient treatment than earlier cohorts. We account for this development in all empirical models (see the method section).

[2] In Denmark, the overall formal assessment of the student at the end of ninth grade consists of both exam results and teacher evaluation grades based on the teachers' assessment of the students' skills. All grades use the same scale (Ministry of Children and Education 2013).

[3] Graduates from public and private school are included, but not home schooled students.

[4] Children diagnosed with ADHD, who do not begin medical treatment for their disorder, are excluded from the sample. Relying on Danish national medical prescription guidelines (Danish Health and Medicines Authority 2014), we expect that excluded cases have only mild symptoms of ADHD.

[5] Excluding these groups from the analysis does not alter significance nor effect sizes of treatment estimates (results not presented but available upon request).

[6] Figures are available upon request.

[7] As we show in the section on data, initial prescribers in our data have undergone thorough medical examination to establish whether they suffer from ADHD and whether they should be treated

medically. Consequently, it is unlikely that they should drop out of medical treatment due to misdiagnosis or mild ADHD symptoms.

[8] Slow-release methylphenidate products and Strattera®--a non-stimulant second line alternative--has been introduced to the Danish market (Pottegård et al. 2012).

[9] However, the fact that the share of DPT is relatively stable across time substantiates the claim that –for unpredictable biological reasons—a proportion of people who do not respond to ADHD medication exist, leaving DPT independent of the relationship between CPT and APT.

[10] We have tested this by allowing birth cohort to affect outcomes in eq. (1) directly over and above that implied by the Heckman selection model (selection into the analysis sample). Including birth cohort in outcome models hardly changes the estimated effects of DPT and APT.

[11] Students diagnosed and treated just after graduating ninth grade seem to have lower grades than students diagnosed before or more than a year later than graduation. This indicates that students diagnosed with ADHD immediately upon graduating ninth grade may initiate treatment due to low school-leaving GPA's. Therefore, we include in the placebo sample those students who initiate treatment more than two years after their ninth grade exams.

[12] Additional controls included in the analysis but not shown are noted below table 3. Results are available upon request.

REFERENCES

Adler, N.E. and K. Newman. 2002. "Socioeconomic disparities in health: pathways and policies." *Health affairs* 21(2):60-76.

Aizer, A. 2008. "Peer Effects and Human Capital Accumulation: the Externalities of ADD.". *NBER* Working Paper 14354.

Alwin, D.F. and R.M. Hauser. 1975. "The decomposition of effects in path analysis." *American Sociological Review*:37-47.

Angrist, J. and G. Imbens. 1994. "Identification and estimation of local average treatment effects." *Econometrica* 62(2):467-475.

Angrist, J.D. and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Massachusetts.

Barkley, R.A. 1977. "A Review Of Stimulant Drug Research With Hyperactive Children." *Journal of Child Psychology and Psychiatry* 18(2):137-165.

Baumgardner, D., A. Schreiber, J. Havlena, F. Bridgewater, D. Steber, and M. Lemke. 2010. "Geographic Analysis of Diagnosis of Attention-Deficit/Hyperactivity Disorder in Children: Eastern Wisconsin, USA." *International Journal of Psychiatry in Medicine* 40(4).

Biederman, J., M.C. Monuteaux, A.E. Doyle, L.J. Seidman, T.E. Wilens, F. Ferrero, C.L. Morgan, and S.V. Faraone. 2004. "Impact of Executive Function Deficits and Attention-Deficit/Hyperactivity Disorder (ADHD) on Academic Outcomes in Children." *Journal of Consulting and Clinical Psychology* 72(5):757-766.

Boyle, M.H. and A.R. Jadad. 1999. "Lessons From Large Trials: The MTA Study as a Model for Evaluating the Treatment of Childhood Psychiatric Disorder." *Can J Psychiatry* 44(10):991-998.

Breslau, J., E. Miller, N. Breslau, K. Bohnert, V. Lucia, and J. Schweitzer. 2009. "The Impact of Early Behavior Disturbances on Academic Achievement in High School." *Pediatrics* 123(6):1472-1476.

Bruckner, T.A., A. Hodgson, C.B. Mahoney, B.D. Fulton, P. Levine, and R.M. Scheffler. 2012. "Health care supply and county-level variation in attention-deficit hyperactivity disorder prescription medications." *Pharmacoepidemiology and Drug Safety* 21(4):442-449.

Carlson, C.L. and M.R. Bunner. 1993. "Effects of Methylphenidate on the Academic Performance of Children with Attention-Deficit Hyperactivity Disorder and Learning Disabilities." *School Psychology Review* 22(2):184-198.

Connor, D.F. 2002. "Preschool attention deficit hyperactivity disorder: a review of prevalence, diagnosis, neurobiology, and stimulant treatment." *Journal of Developmental & Behavioral Pediatrics* 23(1):S1.

Conrad, P. 2006. *Identifying hyperactive children. The medicalization of deviant behavior. The classic study of ADHD*. Burlington: Ashgate.

Conrad, P. and J.W. Schneider. 1992. *Deviance and Medicalization. From Badness to Sickness*. Philadelphia: Temple University Press.

Conti, G., J. Heckman, and S. Urzua. 2010. "The Education-Health Gradient." *The American Economic Review* 100(2):234-238.

Cuffe, S., R. McKeown, K. Jackson, C. Addy, A. R., and C. Garrison. 2001. "Prevalence of attention-deficit/hyperactivity disorder in a community sample of older adolescents." *J Am Acad Child Adolesc Psychiatry* 40(9):1037-1044.

Currie, J. and M.B. Stabile. 2004. "Child Mental Health and Human Capital Accumulation: The Case of ADHD." *On-Line Working Paper Series, California center for Population Research, UC Los Angeles*:1-42.

Currie, J., M.B. Stabile, and L.E. Jones. 2013. "Do Stimulant Medications Improve Educational and Behavioral Outcomes for Children with ADHD?": National Bureau of Economic Research.

Dalsgaard, S., H.S. Nielsen, and M. Simonsen. 2012. "The Effects of Pharmacological Treatment of ADHD on Children's Health." in *Discussion paper series / Forschungsinstitut zur Zukunft der Arbeit, 6714*, edited by IZA. Bonn.

—. 2014. "Consequences of ADHD medication use for children's outcomes." *Journal of Health Economics* 37:137–151.

Danish Health and Medicines Authority. 2010. "Mere end ti-dobling i antallet af personer i behandling med medicin mod ADHD på ti år." Pp. 1-13. Danish Health and Medicines Authority, report.

Danish Health and Medicines Authority. 2014. "National klinisk retningslinje for udredning og behandling af ADHD hos børn og unge." Danish Health and Medicines Authority.

Duncan, G.J., C.J. Dowsett, A. Claessens, K. Magnuson, A.C. Huston, P. Klebanov, L.S. Pagani, L. Feinstein, M. Engel, J. Brooks-Gunn, H. Sexton, K. Duckworth, and C. Japel. 2007. "School Readiness and Later Achievement." *Developmental Psychology* 43(6):1428-1446.

Dunlop, S., P.C. Coyte, and W. McIsaac. 2000. "Socioeconomic status and the utilisation of physicians' services: results from the Canadian National Population Health Survey." *Social Science & Medicine* 51(1):123-133.

Education and Culture DG. 2009. "ECTS Users' Guide." Luxembourg: Office for Official Publications of the European Communities.

Eisenberg, D., E. Golberstein, and J.B. Hunt. 2009. "Mental Health and Academic Success in College." *The B.E. Journal of Economic Analysis & Policy* 9(1).

Fallesen, P. and C. Wildeman. Forthcoming. "The Effect of Medical Treatment of Attention Deficit Hyperactivity Disorder (ADHD) on Foster Care Caseloads: Evidence from Danish Registry Data." *Journal of Health and Social Behavior*.

Faraone, S., R. Perlis, A. Doyle, and J. Smoller. 2005. "Molecular genetics of attention-deficit/hyperactivity disorder." *Biological Psychiatry* 57:1313–1323.

Fletcher, J. and S.F. Lehrer. 2009. "Using genetic lotteries within families to examine the causal impact of poor health on academic achievement." *NBER* Working Paper 15148.

Fletcher, J. and B. Wolfe. 2007. "Child mental health and human capital accumulation: The case of ADHD revisited." *Journal of Health Economics* 27(3):794-800.

Heckman, J.J. 1977. "Dummy endogenous variables in a simultaneous equation system." National Bureau of Economic Research Cambridge, Mass., USA.

Hecktman, L., H. Abikoff, R.G. Klein, G. Weiss, C. Respitz, J. Kouri, C. Blum, B. Greenfield, J. Etcovitch, K. Fleiss, and S. Pollack. 2004. "Academic Achievement and Emotional Status of Children With ADHD Treated With Long-Term Methylphenidate and Multimodal Psychosocial Treatment." *Journal of American Academic Child and Adolescence Psychiatry* 43(7):812-819.

Heinonen, K., K. Raikkonen, A.K. Pesonen, S. Andersson, E. Kajantie, J.G. Eriksson, D. Wolke, A. Lano, and A.-K. Pesonen. 2010. "Behavioural symptoms of attention deficit/hyperactivity disorder in preterm and term children born small and appropriate for gestational age: A longitudinal study." *BMC Pediatrics* 10.

Jackson, M. 2009. "Understanding Links Between Adolescent Health and Educational Attainment." *Demography* 46(4):671-694.

—. 2010. "A life course perspective on child health, cognition and occupational skill qualifications in adulthood: evidence from a British cohort." *Social Forces* 89(1):89.

Jæger, M.M. 2011. "Does Cultural Capital Really Affect Academic Achievement? New Evidence from Combined Sibling and Panel Data." *Sociology of Education* 84(4):281-298.

Kessler, R.C., O. Demler, R.G. Frank, M. Olfson, H.A. Pincus, E.E. Walters, P. Wang, K.B. Wells, and A.M. Zaslavsky. 2005. "Prevalence and treatment of mental disorders, 1990 to 2003." *New England Journal of Medicine* 352(24):2515-2523.

Kessler, R.C., C.L. Foster, W.B. Saunders, and P.E. Stang. 1995. "Social consequences of psychiatric disorders, I: Educational attainment." *The American journal of psychiatry* 152(7):1026.

Kessler, R.C., K.A. McGonagle, S. Zhao, C.B. Nelson, M. Hughes, S. Eshleman, H.U. Wittchen, and K.S. Kendler. 1994. "Lifetime and 12-month prevalence of DSM-III- R psychiatric disorders in the United States. Results from the National Comorbidity Survey." *Archives of General Psychiatry* 51(1):8.

Kildemoes, H.W., H.T. Sørensen, and J. Hallas. 2011. "The Danish national prescription registry." *Scandinavian journal of public health* 39(7 suppl):38-41.

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.

Linnet, K.M., K. Wisborg, E. Agerbo, N.J. Secher, P.H. Thomsen, and T.B. Henriksen. 2006. "Gestational age, birth weight, and the risk of hyperkinetic disorder." *Archives of disease in childhood* 91(8):655-660.

Madaan, V., S. Kinnan, j. Daughton, and C.J. Kratochvil. 2006. "Innovations and recent trends in the treatment of ADHD." *Expert review of neurotherapeutics* 6(9):1375-1385.

Mannuzza, S., R.G. Klein, A. Bessler, P. Malloy, and M.E. Hynes. 1997. "Educational and occupational outcome of hyperactive boys grown up." *Journal of the American Academy of*

Child and Adolescent Psychiatry 36(9):1222-1227.

Mannuzza, S., R.G. Klein, A. Bessler, P. Malloy, M. Lapadula, and M. LaPadula. 1993. "Adult outcome of hyperactive boys. Educational achievement, occupational rank, and psychiatric status." *Archives of General Psychiatry* 50(7):565-575.

McLeod, J. and K. Kaiser. 2004. "Childhood Emotional and Behavioral Problems and Educational Attainment." *American Sociological Review* 69(5):636-658.

McLeod, J.D., R. Uemura, and S. Rohrman. 2012. "Adolescent mental health, behavior problems, and academic achievement." *Journal of Health and Social Behavior* 53(4):482-497.

Miech, R.A., A. Caspi, T.E. Moffitt, B.R.E. Wright, and P.A. Silva. 1999. "Low Socioeconomic Status and Mental Disorders: A Longitudinal Study of Selection and Causation during Young Adulthood." *The American Journal of Sociology* 104(4):1096-1131.

Ministry of Children and Education. 2007. "BEK nr 262 af 20/03/2007 Gældende (Bekendtgørelse om karakterskala og anden bedømmelse)." in *262*. Lovtidende A: Ministry of Children and Education

—. 2013. "LBK nr 521 af 27/05/2013 Gældende (Folkeskoleloven)." in *521*. Lovtidende A: Ministry of Children and Education.

Molina, B.S., S.P. Hinshaw, J.M. Swanson, L.E. Arnold, B. Vitiello, P.S. Jensen, J.N. Epstein, B. Hoza, L. Hechtman, and H.B. Abikoff. 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 & Adolescent Psychiatry* 48(5):484-500.

MTA Cooperative Group. 1999. "A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder." *Archives of General Psychiatry* 56(12):1073-1086.

MTA Coorporate Group. 2004. "National Institute of Mental Health Multimodal Treatment Study of ADHD Follow-up: 24-Month Outcomes of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder." *Pediatrics* 113(4):754–761.

Pelham, W.E. and B.H. Smith. 2000. "Prediction and Measurement of Individual Responses to Ritalin by Children and Adolescents with Attention Deficit Hyperactivity Disorder." Pp. 193-218 in *Ritalin, theory and practice*, edited by L.L. Greenhill and B.B. Osman. Larchmont, NY: Mary Ann Liebert Publishers.

Pliszka, S.R. 2003. "Non-stimulant treatment of attention-deficit/hyperactivity disorder." *CNS spectrums* 8(4):253-258.

Polanczyk, G., M.S.d. Lima, B.L. Horta, J. Biederman, and L.A. Rohde. 2007. "The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis." *American Journal of Psychiatry* 164(6):942-948.

Pottegård, A., B. Bjerregaard, D. Glintborg, J. Hallas, S. Moreno, K. Glintborg, D. Hallas, J. Moreno, and I. Moreno. 2012. "The use of medication against attention deficit hyperactivity disorder in Denmark: a drug use study from a national perspective." *European Journal of Clinical Pharmacology* 68(10):1443-1450.

Powell, S.G., P.H. Thomsen, M. Frydenberg, and H. Rasmussen. 2011. "Long-Term Treatment of ADHD with Stimulants: A Large Observational Study of Real-Life Patients." *Journal of attention disorders* 15(6):439-451.

Rodriguez, A., M.-R. Järvelin, C. Obel, A. Taanila, J. Miettunen, I. Moilanen, T.B. Henriksen, K. Pietiläinen, H. Ebeling, and A.J. Kotimaa. 2007. "Do inattention and hyperactivity symptoms equal scholastic impairment? Evidence from three European cohorts." *BMC Public Health* 7(1):327-335.

Royston, P. and D.G. Altman. 1994. "Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling." *Applied Statistics* 43(3):429-467.

Sanchez, R.J., M.L. Crismon, J.C. Barner, T. Bettinger, and J.P. Wilson. 2005. "Assessment of adherence measures with different stimulants among children and adolescents." *Pharmacotherapy* 25(7):909-917.

Schachar, R. and R. Sugarman. 2000. "Combining Parent Training and Medication in the Treatment of Attention Deficit Hyperactivity Disorder." Pp. 87-109 in *Ritalin, theory and practice*, edited by L.L. Greenhill and B.B. Osman. Larchmont, NY: Mary Ann Liebert Publishers.

Schermerhorn, A.C., B.M. D'Onofrio, W.S. Slutske, R.E. Emery, E. Turkheimer, K.P. Harden, A.C. Heath, and N.G. Martin. 2012. "Offspring ADHD as a risk factor for parental marital problems: controls for genetic and environmental confounds." *Twin Research and Human Genetics* 15(6):700-713.

Spencer, T., J. Biederman, T. Wilens, M. Harding, D. O'Donnell, and S. Griffin. 1996. "Pharmacotherapy of Attention-Deficit Hyperactivity Disorder across the Life Cycle."

Journal of the American Academy of Child & Adolescent Psychiatry 35(4):409-432.

Sprague, R.L. and E.K. Sleator. 1977. "Methylphenidate in Hyperkinetic Children: Differences in Dose Effects on Learning and Social Behavior." *Science, New Series* 198(4323):1274-1276.

Stolzenberg, R.M. 1980. "The measurement and decomposition of causal effects in nonlinear and nonadditive models." *Sociological methodology*:459-488.

Swanson, J. 2003. "Compliance with stimulants for attention-deficit/hyperactivity disorder." *CNS Drugs* 17(2):117-131.

UNESCO Institute of Statistics. 2012. "International Standard Classification of Education. ISCED 2011." Montreal, Canada: UNESCO Institute for Statistics.

Vaughan, B.S., J.S. March, and C.J. Kratochvil. 2012. "The evidence-based pharmacological treatment of paediatric ADHD." *International Journal of Neuropsychopharmacology* 15:27-39.

Visser, S., R. Bitsko, M. Danielson, R. Perou, and S. Blumberg. 2010. "Increasing Prevalence of Parent-reported Attention-deficit/Hyperactivity Disorder Among Children - United States, 2003-2007." *Morbidity & Mortality Weekly Report* 59(44):1439-1443.

Wang, J. 2007. "Mental health treatment dropout and its correlates in a general population sample." *Medical Care* 45(3):224-229.

Wilens, T.E. and T.J. Spencer. 2000. "The Stimulants Revisited." *Child and adolescent psychiatric clinics of North America* 9(3):573-603.

Table 1. Descriptive statistics for all background variables. Population sample of compulsory school graduates from years 2002-2011 and the ADHD sample.

	Population sample			ADHD sample					
				Treatment sample			Placebo sample		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Individual characteristics									
Male	0.50	0.50	571107	0.74*	0.44	3738	0.58*	0.49	3785
Gestational age	39.58	1.80	529033	39.32*	2.04	3495	39.52	1.88	3545
Birth weight	3460.78	567.00	531476	3456.23	620.17	3515	3386.56*	582.17	3562
Immigrant status									
Non-immigrant (reference)	0.91	0.29	571072	0.96*	0.21	3738	0.96*	0.20	3785
Immigrant	0.04	0.21	571072	0.02*	0.15	3738	0.03*	0.16	3785
Descendant of immigrant	0.05	0.22	571072	0.02*	0.15	3738	0.02*	0.13	3785
Graded with the new rating scale (from 2008)	0.42	0.49	571107	0.75*	0.43	3738	0.08*	0.27	3785
Birth cohort[a]	1990.53	2.84	571107	1992.32*	2.32	3738	1988.54*	1.91	3785
Age at treatment start				13.07	2.81	3738	20.93	1.93	3785
Treatment pattern									
DPT				0.11	0.31	3738	0.32	0.47	3785
APT				0.74	0.44	3738	0.38	0.49	3785
CPT				0.15	0.36	3738	0.30	0.46	3785
Residential region at age three									
Capital (reference)	0.27	0.45	542025	0.32*	0.47	3613	0.23*	0.42	3666
Zealand	0.14	0.35	542025	0.12*	0.33	3613	0.13*	0.34	3666
South Denmark	0.23	0.42	542025	0.17*	0.38	3613	0.17*	0.38	3666
Central Jutland	0.24	0.42	542025	0.27*	0.44	3613	0.34*	0.48	3666
Northern Jutland	0.11	0.31	542025	0.10*	0.30	3613	0.11	0.32	3666
Bornholm	0.01	0.09	542025	0.01*	0.12	3613	0.01	0.10	3666
Residential region at graduation age									
Capital (reference)	0.25	0.43	570740	0.27*	0.44	3738	0.19*	0.40	3784
Zealand	0.16	0.37	570740	0.15	0.36	3738	0.16	0.36	3784
South Denmark	0.23	0.42	570740	0.18*	0.38	3738	0.17*	0.37	3784
Central Jutland	0.24	0.42	570740	0.28*	0.45	3738	0.35*	0.48	3784
Northern Jutland	0.11	0.32	570740	0.10*	0.30	3738	0.12	0.33	3784
Bornholm	0.01	0.11	570740	0.02*	0.14	3738	0.01	0.10	3784

Parent characteristics

Mother diagnosed with ADHD	0.01	0.07	571107	0.09*	0.28	3738	0.07*	0.25	3785
Father diagnosed with ADHD	0.00	0.06	571107	0.05*	0.21	3738	0.03*	0.18	3785
Mother's age at childbirth	28.57	4.77	568662	28.23*	4.87	3733	27.19*	4.90	3779
Father's age at childbirth	31.53	5.68	520182	31.08*	5.78	3276	30.40*	5.86	3324
Mother's education[b]									
No education (reference)	0.14	0.35	571107	0.08*	0.26	3738	0.12*	0.32	3785
ISCED 2	0.17	0.38	571107	0.23*	0.42	3738	0.31*	0.46	3785
ISCED 3A	0.08	0.28	571107	0.09	0.29	3738	0.07*	0.25	3785
ISCED 3C	0.38	0.48	571107	0.42*	0.49	3738	0.34*	0.47	3785
ISCED 5B	0.19	0.39	571107	0.15*	0.36	3738	0.13*	0.34	3785
ISCED 5AC	0.04	0.20	571107	0.03*	0.18	3738	0.02*	0.15	3785
ISCED 6	0.00	0.02	571107	NA	NA	3738	0.00	0.02	3785
Father's education [b]									
No education (reference)	0.26	0.44	571107	0.24*	0.43	3738	0.30*	0.46	3785
ISCED 2	0.11	0.32	571107	0.16*	0.37	3738	0.16*	0.37	3785
ISCED 3A	0.05	0.22	571107	0.04*	0.20	3738	0.04*	0.20	3785
ISCED 3C	0.40	0.49	571107	0.41	0.49	3738	0.37*	0.48	3785
ISCED 5B	0.11	0.31	571107	0.09*	0.29	3738	0.08*	0.27	3785
ISCED 5AC	0.07	0.25	571107	0.06*	0.23	3738	0.05*	0.21	3785
ISCED 6	0.00	0.05	571107	0.00	0.05	3738	0.00	0.04	3785
Mother's yearly income (USD, 2009 prices)[b]	40309	21140	541964	39859*	17353	3607	36501*	16723	3688
Father's yearly income (USD, 2009 prices)[b]	58432	53511	501487	55771*	28338	3213	54118*	36316	3264
Mother has negative income[b]	0.00	0.03	571107	NA	NA	3738	0.00	0.03	3785
Father has negative income[b]	0.00	0.05	571107	0.00	0.037	3738	0.00	0.04	3785
Mother's unemployment rate[b]	138.85	253.76	535712	170.46*	272.29	3607	177.98*	272.42	3647
Father's unemployment rate[b]	75.77	191.92	496443	97.93*	216.60	3213	103.14*	215.03	3231

Note: * indicates that the mean of the ADHD sample is significantly different from the population sample mean at the 5 percent level of signification (estimated from t-tests). NA denotes zero responses for the variable in question. Un-weighted data. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

[a] The analyses use dummies for birth cohort.

[b] Information on parent educational attainment, income, and unemployment is measured one year prior to the birth of the respondent in order to avoid post treatment confounding.

Table 2. Descriptive statistics for unstandardized and standardized dependent variables. Population sample of compulsory school graduates from years 2002-2011 and the ADHD sample.

	Population sample			ADHD sample					
				Treatment sample			Placebo sample		
	Mean [min; max]	SD	N	Mean [min; max]	SD	N	Mean [min; max]	SD	N
Unstandardized variables									
Exam GPA	6.23 [-3.00;12.00]	2.22	571107	4.86* [-3.00;12.00]	2.11	3738	4.71* [-2.00;11.43]	1.98	3785
Teacher evaluation	6.34 [-3.00;12.00]	2.22	571107	4.79* [-3.00;12.00]	2.07	3738	4.64* [-1.00;11.42]	2.00	3785
Standardized variables									
Exam GPA	0.00 [-4.16;2.61]	1.00	571107	-0.62* [-4.16;2.61]	0.95	3738	-0.69* [-3.71;2-35]	0.90	3785
Teacher evaluation	0.00 [-4.20;2.55]	1.00	571107	-0.70* [-4.20;2.55]	0.93	3738	-0.77* [-3.30;2.28]	0.90	3785

Note: * indicates that the mean of the ADHD sample is significantly different from the population sample mean at the 5 percent level of signification (estimated from t-tests). Un-weighted data. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

Table 3. Estimation results for treatment variables and select control variables[a]. OLS regression models and Heckman Two-Stage sample selection model. Outcome is standardized exam GPA. Treatment and placebo ADHD sample.

	Treatment sample			Placebo sample
	OLS no controls	OLS with controls	Two-Stage with controls[b]	OLS with controls
DPT	-0.14*	-0.19**	-0.18**	-0.05
	(0.06)	(0.06)	(0.06)	(0.03)
APT	-0.10*	-0.13**	-0.13**	-0.04
	(0.04)	(0.04)	(0.04)	(0.03)
Male		-0.20***	-0.29***	-0.24***
		(0.03)	(0.07)	(0.03)
Immigrant		-0.28*	-0.26	-0.33**
		(0.14)	(0.14)	(0.12)
Descendant of immigrant		0.02	0.09	0.05
		(0.11)	(0.12)	(0.11)
Mother diagnosed with ADHD		0.12*	-0.14	0.21***
		(0.05)	(0.18)	(0.05)
Father diagnosed with ADHD		0.07	-0.13	0.16*
		(0.07)	(0.15)	(0.08)
Mother's education ISCED 2		-0.00	-0.01	-0.08
		(0.07)	(0.07)	(0.05)
Mother's education ISCED 3A		0.62***	0.61***	0.52***
		(0.08)	(0.08)	(0.07)
Mother's education ISCED 3C		0.13	0.13	0.10
		(0.07)	(0.07)	(0.05)
Mother's education ISCED 5B		0.35***	0.35***	0.42***
		(0.08)	(0.08)	(0.06)
Mother's education ISCED 5AC		0.76***	0.77***	0.67***
		(0.11)	(0.11)	(0.11)
Father's education ISCED 2		0.01	0.00	-0.05
		(0.06)	(0.06)	(0.05)
Father's education ISCED 3A		0.48***	0.52***	0.38***
		(0.09)	(0.09)	(0.08)
Father's education ISCED 3C		0.06	0.07	0.04
		(0.06)	(0.06)	(0.05)
Father's education ISCED 5B		0.28***	0.30***	0.20**
		(0.07)	(0.07)	(0.06)
Father's education ISCED 5AC		0.38***	0.40***	0.49***
		(0.08)	(0.09)	(0.08)
Father's education ISCED 6		-0.15	-0.20	1.00**
		(0.31)	(0.31)	(0.37)
Constant	-0.53***	-0.01	0.57	-1.05**
	(0.04)	(0.37)	(0.53)	(0.38)
Lambda			-0.24	
			(0.16)	
Rho			-0.26	
R^2	0.00	0.12	-	0.17
Observations	3738	3738	567322	3785
Uncensored observations	-	-	3738	-

Note: * p<0.05, ** p<0.01, *** p<0.001. Standard errors reported in parentheses. DPT: Discontinued Pharmacological Treatment, APT: Ambiguous Pharmacological Treatment, CPT: Continuous Pharmacological Treatment (reference).

Outcome variables are standardized within the total population sample. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

[a] Models with controls also include parent age at childbirth, unemployment rate, income, negative income, as well as individual level information on birth weight, gestational age, age at treatment start, and being graded with the new grading scale. Dummies for missing information on all relevant variables are also included. The Two-stage outcome model also includes residential region at graduation age. Full model estimates are available upon request.

[b] Results from the Heckman two-stage selection model estimates of birth cohort and residential region at age three available upon request.

Table 4. Estimation results for treatment variables and select control variables[a]. OLS regression models and Heckman Two-Stage sample selection model. Outcome is standardized teacher evaluation. Treatment and placebo sample.

	Treatment sample			Placebo sample
	OLS no controls	OLS with controls	Two-Stage with controls[b]	OLS with controls
DPT	-0.14*	-0.19**	-0.19**	-0.06
	(0.06)	(0.06)	(0.06)	(0.03)
APT	-0.08	-0.12**	-0.12**	-0.01
	(0.04)	(0.04)	(0.04)	(0.03)
Male		-0.24***	-0.36***	-0.36***
		(0.03)	(0.06)	(0.03)
Immigrant		-0.22	-0.18	-0.25*
		(0.13)	(0.14)	(0.12)
Descendant of immigrant		0.03	0.13	0.10
		(0.11)	(0.12)	(0.11)
Mother diagnosed with ADHD		0.15**	-0.21	0.19***
		(0.05)	(0.18)	(0.05)
Father diagnosed with ADHD		0.05	-0.23	0.16*
		(0.07)	(0.15)	(0.08)
Mother’s education ISCED 2		0.06	0.04	-0.07
		(0.07)	(0.07)	(0.05)
Mother’s education ISCED 3A		0.54***	0.53***	0.49***
		(0.08)	(0.08)	(0.07)
Mother’s education ISCED 3C		0.16*	0.15*	0.11*
		(0.07)	(0.07)	(0.05)
Mother’s education ISCED 5B		0.33***	0.33***	0.43***
		(0.08)	(0.08)	(0.06)
Mother’s education ISCED 5AC		0.70***	0.71***	0.68***
		(0.11)	(0.11)	(0.11)
Father’s education ISCED 2		-0.03	-0.04	-0.06
		(0.06)	(0.06)	(0.05)
Father’s education ISCED 3A		0.40***	0.46***	0.35***
		(0.09)	(0.09)	(0.08)
Father’s education ISCED 3C		0.04	0.06	0.05
		(0.06)	(0.06)	(0.05)
Father’s education ISCED 5B		0.29***	0.31***	0.19**
		(0.07)	(0.07)	(0.06)
Father’s education ISCED 5AC		0.44***	0.47***	0.45***
		(0.08)	(0.08)	(0.08)
Father’s education ISCED 6		0.06	0.04	1.08**
		(0.31)	(0.31)	(0.37)
Constant	-0.63***	0.03	0.84	-1.56***
	(0.04)	(0.36)	(0.52)	(0.38)
Lambda			-0.33	
			(0.16)	
Rho			-0.36	
R^2	0.00	0.11	-	0.18
Observations	3738	3738	567322	3785
Uncensored observations	-	-	3738	-

Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Standard errors reported in parentheses. DPT: Discontinued Pharmacological Treatment, APT: Ambiguous Pharmacological Treatment, CPT: Continuous Pharmacological Treatment (reference).

Outcome variables are standardized within the total population sample. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

[a] Models with controls also include parent age at childbirth, unemployment rate, income, negative income, as well as individual level information on birth weight, gestational age, age at treatment start, and being graded with the new grading scale. Dummies for missing information on all relevant variables are also included. The Two-stage outcome model also includes residential region at graduation age. Full model estimates are available upon request.

[b] Appendix A shows Heckman two-stage selection model estimates of birth cohort and residential region at age three.

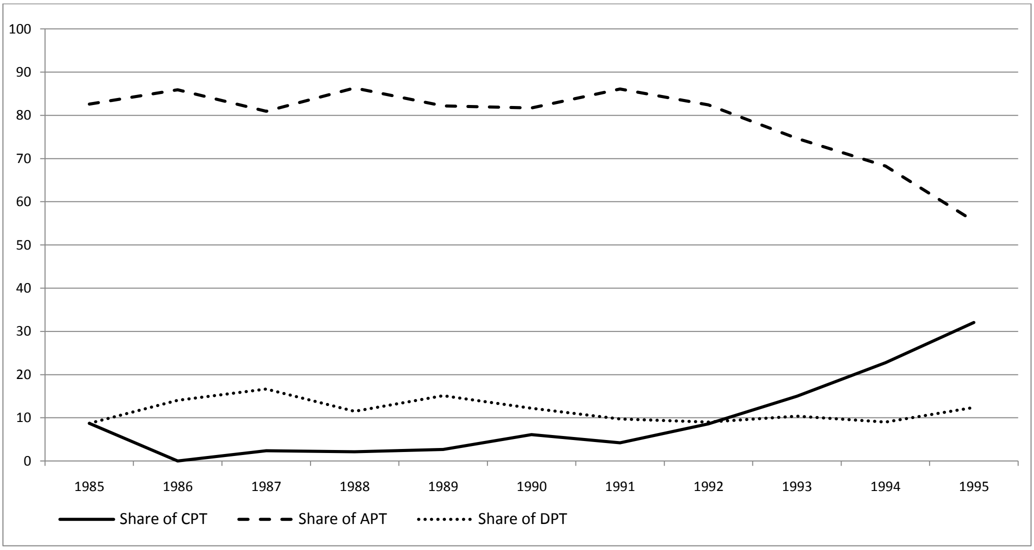

Figure 1. Shares of treatment groups DPT, APT and CPT per birth cohort within the treatment sample (*N* = 3738). Percentages.

Note: Shares sum to a 100 percent per birth cohort. DPT: Discontinued Pharmacological Treatment, APT: Ambiguous Pharmacological Treatment, CPT: Continuous Pharmacological Treatment.

APPENDIX. ADDITIONAL TABLES

Table A1 presents results of an OLS regression model showing that the observed negative effect of having an ADHD diagnosis is the same in the treatment sample as in the placebo sample, when controlling for all background characteristics. Thus, despite some observed differences between students in the treatment sample and students in the placebo sample, they are equally negatively affected by ADHD on our outcome. The placebo sample therefore seems well suited as a comparison to the treatment sample.

Table A1. Effects of having an ADHD diagnosis on standardized exam GPA and teacher evaluation. OLS models including control variables[a]. Population sample of compulsory school graduates from years 2002-2011.

	Exam GPA		Teacher evaluation	
ADHD diagnosis	-0.55***		-0.61***	
(dummy for being in the *treatment* sample)	(0.02)		(0.02)	
ADHD diagnosis		-0.49***		-0.55***
(dummy for being in the *placebo* sample)		(0.01)		(0.01)
Constant	0.98***	0.97***	-0.76***	0.75***
	(0.05)	(0.05)	(0.05)	(0.05)
R^2	0.23	0.23	0.23	0.23
Observations	567322	567369	567322	567369

Note: * p<0.05, ** p<0.01, *** p<0.001. Standard errors reported in parentheses below estimates. Outcome variables are standardized within total population sample. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

[a] Models include controls for all Table 1 variables, except age at treatment start and residential region at age three. Dummies for missing information on all relevant variables are also included. Full model estimates are available upon request.

Table A2 presents additional estimates from selection models, thus supplementing the information shown in tables 3 and 4.

Table A2. Estimation results for residential region at age three and birth cohort dummies from Heckman Two-Stage Selection model. Treatment sample.

Residential region at age three	
Zealand	-0.15***
	(0.02)
South Denmark	-0.18***
	(0.02)
Central Jutland	-0.04*
	(0.02)
Northern Jutland	-0.14***
	(0.02)
Bornholm	0.15*
	(0.06)
Birth cohort	
1985	-0.08
	(0.26)
1986	-0.44
	(0.26)
1987	-0.37
	(0.25)
1988	-0.21
	(0.25)
1989	-0.13
	(0.25)
1990	-0.05
	(0.25)
1991	-0.05
	(0.25)
1992	-0.30
	(0.26)
1993	-0.20
	(0.26)
1994	-0.12
	(0.26)
1995	-0.12
	(0.26)
1996	-0.54
	(0.31)
Observations	567322
Uncensored observations	3738

Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Standard errors reported in parentheses. Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. The Two-stage outcome models for teacher evaluation or exam GPA include controls for all Table 1 variables and dummies for missing information on all relevant variables.

Table A3 shows results for a regression of treatment status (reference is placebo status) on our observables.

Table A3. Regression coefficients estimating the likelihood of being in the treatment sample (reference is placebo sample). Logit models including control variables.

	Full model[a]	Reduced model
Male	0.91***	
	(0.07)	
Birth weight	0.00	
	(0.00)	
Birth weight, missing	1.00*	
	(0.49)	
Gestational age	-0.07**	
	(0.02)	
Gestational age, missing	-3.88***	
	(0.93)	
Graded with new grading scale	1.98***	
	(0.18)	
Immigrant	0.46	0.20
	(0.32)	(0.26)
Descendent of immigrant	0.43	0.43
	(0.27)	(0.27)
Mother diagnosed with ADHD	0.35**	0.28*
	(0.13)	(0.13)
Father diagnosed with ADHD	0.35*	0.33
	(0.18)	(0.17)
Mother’s yearly income	0.00	0.00*
	(0.00)	(0.00)
Father’s yearly income	-0.00	-0.00
	(0.00)	(0.00)
Mother’s income, missing	-0.56	-0.44
	(0.40)	(0.37)
Father’s income, missing	NA	NA
	NA	NA
Mother has negative income	NA	NA
	NA	NA
Father has negative income	-0.35	-0.02
	(0.74)	(0.86)
Mother’s unemployment rate	-0.00	-0.00
	(0.00)	(0.00)
Father’s unemployment rate	-0.00*	-0.00*
	(0.00)	(0.00)
Mother’s unemployment rate, missing	1.06**	0.74*
	(0.37)	(0.30)
Father’s unemployment rate, missing	0.70*	-0.64*
	(0.33)	(0.30)
Mother’s education ISCED 2	0.15	0.13
	(0.15)	(0.14)
Mother’s education ISCED 3A	0.39*	0.34*
	(0.18)	(0.17)
Mother’s education ISCED 3C	0.23	0.20
	(0.15)	(0.14)
Mother’s education ISCED 5B	0.19	0.14
	(0.16)	(0.16)

Mother's education ISCED 5AC	0.13	0.06
	(0.27)	(0.25)
Mother's education ISCED 6	NA	NA
	NA	NA
Father's education ISCED 2	0.01	-0.02
	(0.14)	(0.13)
Father's education ISCED 3A	-0.22	-0.22
	(0.21)	(0.20)
Father's education ISCED 3C	-0.02	-0.06
	(0.12)	(0.11)
Father's education ISCED 5B	0.07	0.01
	(0.16)	(0.15)
Father's education ISCED 5AC	0.02	0.01
	(0.18)	(0.18)
Father's education ISCED 6	0.27	0.19
	(0.65)	(0.68)
Mother's age at childbirth	0.02	0.02
	(0.01)	(0.01)
Mother's age at childbirth, missing	0.97	0.79
	(1.14)	(1.05)
Father's age at childbirth	-0.00	-0.00
	(0.01)	(0.01)
Father's age at childbirth, missing	0.67	0.65
	(0.42)	(0.39)
Birth cohort		
1985	-4.39***	-6.08***
	(0.54)	(0.48)
1986	-5.08***	-6.86***
	(0.49)	(0.43)
1987	-4.92***	-6.69***
	(0.49)	(0.43)
1988	-4.53***	-6.28***
	(0.48)	(0.42)
1989	-4.18***	-5.95***
	(0.48)	(0.42)
1990	-3.87***	-5.58***
	(0.47)	(0.42)
1991	-3.62***	-4.81***
	(0.45)	(0.42)
1992	-3.83***	-3.73***
	(0.40)	(0.42)
1993	-2.04***	-1.92***
	(0.42)	(0.44)
1994	0.36	0.48
	(0.62)	(0.65)
Residential region at graduation		
Zealand	-0.55***	-0.47***
	(0.11)	(0.11)
South Denmark	-0.06	0.01
	(0.11)	(0.10)
Central Jutland	-0.47***	-0.42***
	(0.09)	(0.09)
Northern Jutland	-0.41**	-0.38**
	(0.13)	(0.12)
Bornholm	NA	0.49
	NA	(0.27)
Constant	4.43***	4.34***
	(0.94)	(0.56)

Pseudo R^2	0.47	0.43
Log likelihood	-2783.9	-2974.2

Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Standard errors reported in parentheses. NA denotes zero responses for the variable in question. Sample consists of both placebo and treatment sample ($N = 7522$). Treatment sample: Children medically treated for ADHD *prior to* compulsory school graduation. Placebo sample: Children medically treated for ADHD *after* compulsory school graduation.

[a] Full model includes controls for all Table 1 variables, except age at treatment start and residential region at age three.