Teen antidepressant use and academic achievement*

Sonia Bhalotra
University of Essex and IZA

N. Meltem Daysal
University of Copenhagen and IZA

Nis Vestergård Lydiksen
University of Southern Denmark and VIVE

Mircea Trandafir
University of Southern Denmark and IZA

February 2021

***** Preliminary and incomplete, please do not cite *****

Abstract

We investigate the effects of antidepressant use during adolescence on the educational outcomes of treated children. Using the propensity of the first treating specialist to prescribe antidepressants to other children as instrument, we find large and significant benefits from treatment with antidepressants on test scores, especially math. Although imprecise, our findings suggest girls benefit more than boys from the pharmaceutical treatment of emotional disorders.

Keywords: Antidepressants, education, human capital

JEL Classifications: I11, I12, I18, J13

*Bhalotra: Department of Economics, University of Essex, and IZA (email: srbhal@essex.ac.uk); Daysal: Department of Economics, University of Copenhagen, and IZA (email: meltem.daysal@sam.sdu.dk); Lydiksen: Department of Business and Economics, University of Southern Denmark, and VIVE (email: nisvl@sam.sdu.dk); Trandafir: Department of Business and Economics, University of Southern Denmark, and IZA (email: mircea.trandafir@sam.sdu.dk).
James Fenske, Torben Heien Nielsen, Niels Skipper, and participants at the workshop "Medical Interventions and Socio-Economic Outcomes" provided helpful comments and discussions. The research was conducted with financial support from the Danish Foundation for Independent Research under grant 4182-00231.

1 Introduction

Mental illness is increasingly recognized as a global health issue. Recent estimates find that the burden of mental, neurological, and substance abuse disorders increased by 41 percent in the last 20 years to 13 percent of the total disease burden (WHO, 2018). The most common mental disorders belong to two main diagnostic categories, depressive disorders and anxiety disorders: it is estimated that in 2015 about 4.4 percent of the global population lived with depression and about 3.6 percent with anxiety, which represent an increase of 18.4 and 14.9 percent, respectively, since 2005 (WHO, 2017). Globally, depressive disorders and anxiety disorders rank as the first and sixth largest contributors to non-fatal health loss, respectively. Of particular concern is the fact that the incidence of mental health problems during childhood and adolescence is increasing: depression is one of the leading causes of illness and disability among adolescents, and suicide is one of the leading causes of death among children 15-19 years old (WHO, 2018). Combined with the fact that the median age of onset of all mental disorders is before the age of 14, these imply that there is a need for better understanding the consequences of mental disorders and their treatment in adolescence.

The most commonly used drugs to treat both depression and anxiety are antidepressants (AD). Previous clinical research indicates that most types of AD successfully alleviate short-term negative symptoms of depression, especially among adults (Amsterdam, 1998; Ryan, 2003; Nemeroff, 2007). However, there is no general consensus on their benefits for children and adolescents, including in Denmark. For example, in a letter to the editor of Politiken (January 6th, 2014), Peter Gøtzsche, the then-leader of the Nordic Cochrane Institute, criticizes what he sees as a dangerous over-prescription of drugs by psychiatrists. In their response (Politiken, January 17th, 2014), three of Denmark's most influential psychiatrists (Lars Kessing, Merete Nordentoft and Thomas Middelboe) defend current practices on the basis of the evidence in the literature. The debate among (Danish) psychiatrists on whether the level of use of AD is appropriate, and especially if children and adolescents have any benefits from treatment, still continues. In order to make informed policy decisions, more evidence is needed on the consequences of AD treatment beyond purely clinical outcomes.

In this paper we aim to fill this gap by studying the impact of treatment with AD on the educational outcomes of children. We rely on rich register data from Denmark which allows us to follow children over time and to observe their AD use as well as their academic achievement. Empirical identification of the effects of AD use is complicated by the endogeneity in drug use. For example, higher-achieving children may choose pharmaceutical treatment, leading to a spurious correlation between AD use and education outcomes in simple regressions. To address this issue, we exploit the plausibly exogenous variation in AD use stemming from physician differences in the propensity to prescribe AD. Focusing on children aged 10–15 when first diagnosed, our instrumental variable estimates point to large and significant benefits of AD use on test scores, particularly in math: using

AD leads to an increase in math test scores of 1.162 standard deviations, and moving a child from the 25th to the 75th percentile in the distribution of daily-defined doses increases the math test score by 0.616 standard deviations. The estimated effects have relatively wide confidence intervals, with lower bound as low as 0.045 standard deviations for the effect of ever using AD. For comparison, Norwegian children who received medical treatments soon after birth because of very low birth weight experienced higher test scores by 0.476 standard deviations (Bharadwaj et al., 2013), while Figlio et al. (2014) find that a 10 percent increase in birth weight leads to an increase in test scores of 0.044 standard deviations. We also provide some evidence suggesting that girls benefit more than boys from treatment with AD, and we show that or results are not driven by selective test-taking.

Our study is closely related to previous research in economics examining the effects of mental disorders, particularly depression. Within this literature the majority of papers focuses on how depression affects an individual's labor market outcomes (e.g., Ettner et al., 1997; Berndt et al., 1998; Alexandre and French, 2001; Marcotte and Wilcox-Gök, 2003; Chatterji et al., 2011; Fletcher, 2013) and human capital accumulation (e.g., Kessler et al., 1995; Currie and Stabile, 2009; Fletcher, 2010; Fletcher and Lehrer, 2011; Busch et al., 2014). The impact of depression on other individual outcomes is largely unexplored, one exception being Anderson et al. (2015), who examine the relationship between youth depression and future crime.

Our project is most related to the body of research studying the effects of pharmaceutical treatment of depression. This literature almost exclusively examines how AD use impacts an individual's own health outcomes. Findings from previous clinical trials generally suggest that antidepressants, especially selective serotonin reuptake inhibitors (SSRIs), are successful in alleviating symptoms of depression (e.g., Ryan, 2003; Green, 2003). Recent cross-country evidence from economics also suggests that higher SSRI sales are associated with reduced suicides (Ludwig and Marcotte, 2005; Ludwig et al., 2009). Despite this evidence, there is very limited research on how AD impact the non-health outcomes of adults (Schoenbaum et al., 2002, and Galárraga et al., 2010, on labor market outcomes, Marcotte and Markowitz, 2011, on violent crime) and none that we are aware of for the non-health outcomes of children.

Our paper makes several contributions. First, we examine how treated children are affected by AD use along a set of dimensions that, to the best of our knowledge, were not previously examined. Second, the richness of the administrative data allows us to separately investigate short-term and long-term effects as well as the dynamics over the life-cycle. For example, we can investigate the effects on human capital accumulation and criminal behavior at various points in the life cycle. While some studies do look at the long-term effects of depression, there is no investigation of the relationship between AD and long-term outcomes or how this relationship changes over the life-cycle. Third, we investigate differences in effects by observable individual and family characteristics such as gender and parental income and education. This is especially relevant to understanding the existing health disparities and the intergenerational transmission of health and education. Fourth, our sample includes the entire population as opposed to the previous literature relying on select subpopulations. Finally, our empirical strategy improves upon previous studies by relying on a plausibly exogenous source of variation in antidepressant use.

2 Background

Adolescent mental health The formative years between 10-18 are a crucial time for development and school choices impacting long into adult life. Adolescent mental health problems are common, with WHO estimating 16 percent of the global burden of disease coming from mental health and siucide being the third leading factor for the age group, with depression and anxiety being the primary illnesses.[1] In the Danish setting of our study hospital diagnosis of 10-19 year old's mental health issues tripled in the 200-2013 period Skovlund et al. (2017). Mental health is more difficult to define and measure than somatic diseases with issues ranging from mild, periodic problems, to long-term very

Organization of mental health care Health care in Denmark, including mental health care, is primarily publicly organized and funded. Care is provided trough decentralized health care regions that runs hospitals, and contracts privately practicing physicians in primary care. Care is generally provided free of charge to the patient with some exceptions (Pedersen et al., 2005). General practitioners (GPs), child psychiatrist and pharmacies are the main actors in our analysis and their role in treating depression in children is detailed below.

General practitioners. All patients are assigned a GP as the main entry to mental health care. GP's screens patients and acts as gatekeepers to more specialized treatment. Patients pay no co-payment for services and GPs are instead reimbursed in a tax financed fee-for-service model. The GP will either initiate treatment or refers patients to a specialist. In the majority of non-severe adult cases the GP will administer treatment after diagnosis (Bauer et al., 2012). Generally the GP decides between "watchful waiting", psychoeducation, prescribing antidepressants or a combination of the two latter. For children the GP's options are limited by strong guidelines on prescription of antidepressant to children. Guidelines establish that antidepressant prescription to children is a specialist task because of the uncertain efficacy and possible adverse effects in the group. Following guidelines, this leaves GPs with the option of non-pharmaceutical treatment or referring to a child psychiatrist.

Child psychiatrists are specialists in treating mental illnesses in children. Consultations are free to patients when referred from the GP. Specialists can be either privately practising (reimbursed in a similar way as GPs) or hospital (salaried) doctors at child psychiatric departments. GPs refers

[1]See WHO Mental Health Action Plan 2013-2020.

patients to both types of specialists primarily determined by availability at the time. In our analysis we treat the two types as substitutes as long waiting times limits the degree to which patients can select either. Patients can seek specialist services at private clinics without a referral but must then pay the full cost themselves. The specialist will verify the GPs diagnosis and treat the patient based on a stepped care approach with psychoeducation and psychotherapy being the first options with introduction of pharmaceutical treatment in more severe cases.[2]. Antidepressants are not exclusively prescribed to treat depression but importantly also anxiety and obsessive compulsive disorder.[3] Child psychiatric services have historically faced a significant overdemand. In our study period waiting times for hospital psychiatry where at times up to 200 days and limited the patients ability to choose a specific provider or type of provider.[4] Lack of available specialist also explains part of the significant GP non-compliance with prescription guidelines observed in our data as GPs resort to own prescriptions when a specialist is not available.

Pharmacies. Prescription medication are not provided by the prescribing physician but at central pharmacies. Patients covers part of the cost of medication with a co-payment decreasing with the yearly aggregrated medicine costs and a lower co-payment for children.[5]

Educational System We study outcomes related to the outcome of education within the Danish education system. In the Danish system education is compulsory from the year the child turns seven until the completion of 9th grade (grades being classified as school years from 0-10 where the 10th grade is optional and attended by roughly half of children completing the 9th grade). Further education is voluntary and can be either academically oriented (three year high-school) or vocational. Children either attend public schools or private alternatives both with heavily regulated curriculum and learning goals. A small number of children are home schooled and these children are still tested in similar ways as other children. Schools are either directly managed or overseen by local municipalities. All municipalities implements some form of psycho-social guidance program that aid parents and health care professionals in detecting student mental health issues but schools can neither treat or refer students to treatment. Final examinations, which is the basis of our main outcome, is compulsory for all children in primary school also children with special needs or who is enrolled in some form of special education.[6] While final examination results are not directly related to further educational options they are permanently recorded and are generally considered high stakes tests.

[2]See the guidelines on treatment of depression, anxiety and obsessive compulsive disorder from the association of child psychiatrists

[3]In our study period an equal number of children are prescribed antidepressant for treating depression and anxiety, see the medicine monitoring report from the Danish Health Authority.

[4]See the reporting of waiting times for child psychiatry from the Danish Health Authority.

[5]Cost to children are reimbursed 60 % for first 1,625 dkk, 75 % up till 3,520 dkk , 85 % up till 23,428 after which 100 % of costs are reimbursed.

[6]Individual exceptions are made but there are no general exemptions criteria based on mental health condition. See the law on mandatory final assessment (folkeskoleloven Act No. 313 of 19 April 2006).

3 Empirical strategy

We are interested in estimating the effect of AD use during teen years on the outcomes of treated children. The equation of interest is:

$$Y_i = \beta_0 + \beta_1 AD_i + \delta \mathbf{X}_i + \epsilon_i, \tag{1}$$

where Y_i is an (academic) outcome of child i, AD_i is a measure of AD use during teen years, and $\mathbf{X}_i$ is a vector of child and family characteristics.

Simple estimates of (1) are likely to yield biased estimates of the parameter of interest β_1 because AD use is likely to be correlated with unobservable characteristics that also affect child outcomes. For example, parents who are focused on the academic performance of their children might "shop around" for a physician who will have the same attitude toward treatment with AD as they do. In order to plausibly estimate the causal effect of AD use, we exploit the plausibly exogenous physician propensity to prescribe AD to *other* children in an instrumental variables (IV) framework. This strategy is similar to the one used in other contexts by, among others, Duggan (2005), Doyle (2007, 2008), Maestas et al. (2013), and Dalsgaard et al. (2014).

The intuition of our instrument is as follows. Consider two identical children, child A and child B, who both suffer from similar mental health issues that can be treated with AD (typically depression or anxiety). These children are both referred by their general practitioners to a specialist. Child A is seen by a physician who tends to prescribe AD for the treatment of mental health issues, but the specialist who sees child B is less likely to do so. As a result, child A will have a higher chance of being prescribed and using AD than child B, so any difference in outcomes between the two will likely be due to differences in (the probability of) the use of AD.

Let PP_i be a measure of the propensity to prescribe AD of the specialist who treats child i. The first stage in our IV setup describes the relationship between teen AD use and physician propensity to prescribe AD:

$$AD_i = \alpha_0 + \alpha_1 PP_i + \delta_{FS} \mathbf{X}_i + u_i, \tag{2}$$

while the reduced form equation relates child outcomes to the propensity to prescribe of the treating specialist:

$$Y_i = \gamma_0 + \gamma_1 PP_i + \delta_{RF} \mathbf{X}_i + v_i. \tag{3}$$

Consistent with the previous literature, we construct the propensity to prescribe measure as the fraction of children receiving AD among those treated by the specialist, excluding child i themselves:

$$PP_i = \frac{\sum_{j \neq i} AD_j \mathbb{1}(CP_j = CP_i)}{\sum_{j \neq i} \mathbb{1}(CP_j = CP_i)}, \tag{4}$$

where $\mathbb{1}(CP_j = CP_i)$ indicates whether children i and j are treated by the same specialist (psychiatrist or child psychiatrist).

In order for our instrument to be valid, it needs to satisfy several conditions. First, it needs to be relevant, meaning that it needs to have an effect on the probability of receiving treatment with AD. This condition can be easily verified using the first-stage regression (2).

Second, the instrument needs to be as-good-as-random. In our context, this implies that a specialist's propensity to prescribe AD should not be correlated with (observable and unobservable) characteristics of the child and their family. One potential scenario which would violate this assumption is when parents who are very invested in their children's academic performance "shop around" for a specialist who is more likely to prescribe AD to their children. However, such a scenario is unlikely to be a main driver in our context. As mentioned in section 2, there is a limited number of specialists operating in Denmark, particularly child psychiatrists. During our sample period, there are only 26 child psychiatrists treating children, meaning that in most cases it would be very difficult for children to change specialists. Still, in order to reduce the potential endogeneity in the treating specialist, we assign to each child the first specialist who treats them and we abstract from any subsequent changes in treating physicians.

Third, the instrument should only affect outcomes through the endogenous variable, i.e., a specialist's propensity to prescribe AD should impact child outcomes only through its effect on the probability that the child uses AD. This assumption precludes cases in which the propensity to prescribe captures some other characteristics of the physician, such as their overall quality, or of their interaction with their patients (e.g., physicians with a lower propensity to prescribe AD might be better at non-pharmaceutical treatment).

Finally, the monotonicity assumption requires that children affected by the instrument should all be impacted in the same direction. In other words, a specialist who is more likely to prescribe AD should not increase the chances that some of their patients are treated with AD while decreasing the chances of others.

All the assumptions other than the relevance assumption are untestable in practice. Therefore, in section 5 we present several pieces of suggestive evidence that can shed more light on their plausibility. If all the assumptions hold, then our IV estimate can be interpreted as an average marginal treatment effect, the average effect of AD use on children who are on the margin to receive it.

4 Data

Our analysis requires information on physician prescription behavior, the use of AD, educational attainment, and background characteristics of children and their families. We obtain this information by combining several Danish administrative databases using unique person and physician practice identifiers.

The *National Prescription Register* records all prescriptions filled at pharmacies in Denmark starting from 1995 (Kildemoes et al., 2011). It includes detailed information on the medication, such as the class according to the Anatomical Therapeutic Chemical (ATC) Classification System, and the amount included in the prescription measured in defined daily doses (DDD).[7] It also records the exact date when the prescription was filled, as well as the personal identifier of the person for whom the prescription was written and the identifier of the prescribing physician. The first identifier allows us to link prescriptions of AD to individuals and to construct our main independent variable, an indicator for whether the person uses AD at any point between the ages of 10 and 15. We use the second identifier to link prescriptions to physicians and hence to construct our instrument, the propensity to prescribe of a given specialist. At this point we should note that we only observe filled prescriptions, not actual consumption of AD or unfilled prescriptions. Therefore, our estimated effect is an intention-to-treat effect combining the effect of physicians prescribing and of patients filling prescriptions, rather than the effect of actually using AD.

The *Health Insurance Register* holds information on reimbursements to private-practice physicians, both general practitioners and specialists, for all patient-related services covered by the national health insurance after 1990 (Andersen et al., 2011). Each record includes the personal identifier of the patient, the identifier and the specialty of the treating physician practice, limited information on the nature of treatment provided (often just a consultation service), and the date of the reimbursement request.[8] Importantly, although this register does not record information on diagnoses, it allows us to identify children who were treated by child psychiatrists, the first specialist who treats them, and the approximate date of that treatment.[9] We use this information to select our sample, but also to determine the set of patients that are used to construct the propensity to prescribe of each specialist.

We obtain educational outcomes from two registers. The *Academic Achievement Register* records the grades for all subjects in the graduating class of primary school (9th grade or 10th grade) and the type of examination (standardized exit test or teacher-assessed regular course exam) for all students graduating in or after 2001 (Jensen and Rasmussen, 2011). From this register we keep the grades in the written part of the exit tests in math and in Danish, which we then standardize at the cohort level to have mean zero and variance one. The second source of education data is the *Population Education Register*, which provides the highest level of completed education and current enrollment

[7]The WHO defines the DDD as "the assumed average maintenance dose per day for a drug used for its main indication in adults." The actual prescribed dose may differ from this as a function of disease severity or patient characteristics such as age, weight, etc.

[8]The date of the reimbursement request is not the same as the date when the treatment was performed. However, it is reasonable to assume that the request is filled relatively soon after the treatment.

[9]The register provides the identifier of the *practice* and not of individual physicians. General practitioner practices, especially in urban areas, typically involve more than one physician. However, specialist practices, especially for (child) psychiatry, tend to be single-physician practices.

for the entire population for each year starting from 1981 (Jensen and Rasmussen, 2011). From this register we extract information on the number of years of schooling obtained by the parents. In a future version, we will also construct indicators for enrollment in education beyond primary school (i.e., beyond compulsory education) and its type (i.e., academic or vocational).

The *National Patient Register* records all inpatient, outpatient and ER visits to Danish hospitals starting in 1977 (Lynge et al., 2011). It provides information on the personal identification number of the patient, the hospital identification number, the date and time of the visit, the date and time of discharge in case of admission, and the main diagnosis recorded according to the International Classification of Diseases (ICD-8 until 1995, ICD-10 afterwards). In a future version of the paper we will use this information to construct variables related to hospital visits for specific reasons (e.g., self-harm, or attempted suicide) and to total length of stay. A related source of data, the *Psychiatric Central Research Register*, keeps track of all inpatient, outpatient and ER visits to the psychiatric departments of Danish hospitals starting from 1977 (Mors et al., 2011). It includes the same type of information as the National Patient Register and can be used to construct variables related to the use of psychiatric services in hospitals.

We use several other national registers to extract background information on children and their families. The *Population Register* provides, among others, information on gender, birth date, marital status, and immigration status for all Danish residents on January 1 of each year starting from 1980. We construct indicators for whether the child is an immigrant (first- or second-generation) or a boy, for the age at first contact with a specialist (which we will call "age at diagnosis" in the rest of the paper), for the year of birth, and for the municipality of residence at the time of diagnosis. The register also lists the personal identifier of the mother and father for virtually all persons born after 1960 (Pedersen, 2011), which allows us to link children to both parents and siblings. For each parent we construct indicators for whether they are an immigrant (first- or second-generation), married, and for their age. In order to avoid any potential reverse causality, all parental characteristics refer to the year when the child turns 6 years old (the school-starting age). The *Register-based Labour Force Statistics* record the labor force participation status of all individuals in the population at the end of November of each year starting from 1980, allowing us to construct indicators for whether each parent is employed when the child turns 6 (Petersson et al., 2011). Finally, the *Income Statistics Register* provides information extracted from tax records on total annual gross income and salaries for all Danish residents beginning in 1980 (Baadsgaard and Quitzau, 2011). As before, we use this to construct indicators for the income decile in the full population in which each parent's total income was situated in the year when the child is 6 years old.[10] Note that, when used on their own, all

[10]We miss parental information for a small number of cases. In these cases, we assign the mode of each variable to the missing value and we include an indicator for missingness. In section 5 we test the robustness of our results to alternative approaches.

monetary variables are deflated to 2015 Danish kroner.

We would like to restrict our analysis to the population at risk of being prescribed antidepressants. As a first step, this would mean restricting our analysis to children diagnosed with an emotional disorder. The first such diagnosis is made by the general practitioner. As mentioned above, the data available from general practitioners, the Health Insurance Register, only include services provided and not diagnoses. Fortunately, as mentioned in section 2, the guidelines require that general practitioners refer children diagnosed with mental disorders to a psychiatrist specialized in treating children, i.e., either a private-practice child psychiatrist or one employed in a hospital psychiatric department. Therefore, we can identify children with mental disorders as children who have a consultation with a child psychiatrist or with a psychiatric department.

In order to eliminate the endogeneity in treating physician due to parents "shopping around" specialists, we assign children to the first specialist treating them. We then construct our instrument as the propensity to prescribe AD of the first treating specialist. This raises two issues. First, we would be able to restrict the sample to children diagnosed with an emotional disorder only for children whose first contact is with a hospital psychiatric department because we only have information on diagnoses from hospital visits. These are generally the more serious cases, so only including these children would not give us a very clear picture of the effect of treatment with AD.[11] In addition, the data available does not allow us to link (physicians employed at) psychiatric departments to prescriptions, meaning that we are not able to construct our instrument for these children.

Instead, we restrict ourselves to children whose first treatment is provided by a private-practice child psychiatrist, and we restrict the sample in a way that presumably get us closer to the population at risk (i.e., diagnosed with an emotional disorder).[12] Specifically, we focus only on children who have their first consultation with a psychiatrist after turning 10 years old.[13] By doing so we likely exclude most children diagnosed with the other major mental disorder diagnosed in childhood, attention-deficit hyperactivity disorder, because about 80 percent of them are diagnosed between the ages of 4 and 11 (Kessler et al., 2007; Dalsgaard et al., 2014). In addition, the incidence of emotional disorders rises sharply in adolescence and especially after puberty (Costello et al., 2005; Thapar et al., 2012; Wesselhoeft et al., 2015; National Health Service, 2018). Finally, emotional disorders are the most common types of mental disorders among adolescents (National Health Service, 2018).

We are interested in the effect of AD use on academic achievement, specifically on the scores in

[11] See, for example, the description of how the psychiatric departments in the Capital Region (one of the five health regions in Denmark) decide on whether to accept referrals from general practitioners for mental disorders, available at https://www.psykiatri-regionh.dk/undersoegelse-og-behandling/Behandling/Center-for-visitation-og-diagnostik/fagperson/Documents/Visitationskriterier.pdf.

[12] A very small number of children are treated by general psychiatrists. We exclude them from the analysis sample and we check whether this affects our main estimates in section 5.

[13] We define a first consultation as the situation when the child had no consultation with a (child) psychiatrist or a psychiatrist after the age of 6, i.e., after the school-starting age.

tests taken at the age of 16 or 17 and subsequent school enrollment. Therefore, we also restrict the sample to children diagnosed by the age of 15. These restrictions imply that our sample consists of children born between 1986 and 1999 who had a consultation with a child psychiatrist at some point between the ages of 10 and 15.[14] This results in a total sample of 5,378 children. We further impose the constraint that the treating specialist should treat at least 10 children aged 10–15 in the year of the first consultation, which results in a final analyses sample of 5,373 children, of which 1,944 eventually use AD before the age of 16 and 3,429 do not.[15]

Table 1 describes the analysis sample, both overall and separately for children who use AD by age 15 and those who do not, and compares it to the full sample of children who use specialized psychiatric services (at a hospital or with a privately-practicing specialist) for the first time between the ages of 10–15. A comparison of columns 1 and 2 shows that our analysis sample is more likely to include girls, native Danes, children from smaller families, and children diagnosed at a younger age (Panel A). Panels B and C show that the parents of children in our analysis sample tend to be slightly older and better educated, and more likely to be natives, married, and to have higher income. In order to summarize these characteristics, we construct predicted outcomes by regressing our measures of educational attainment on all the control variables discussed above in the entire population of children born between 1986–1999. The results, listed in Panel D, suggest that children in our analysis sample are positively selected as compared to the population of 10–15 year old children who received psychiatric treatment, as they are predicted to have higher test scores in both Danish and math.

In the rest of the paper we focus only on our analysis sample, and we start by comparing children who filled an AD prescription between the ages of 10–15 and children who do not. Panel A in the Table shows that children treated with AD are much more likely to be girls, natives, to be diagnosed when they are older, and to come from smaller families. Interestingly, it does not appear that subsequent fertility is affected given that the number of younger siblings is similar across the two groups. In terms of parental characteristics (Panels B and C), there seem to be some statistically significant but small differences between children who are prescribed AD and those who are not. For example, children who fill an AD prescription tend to have parents who are native Danes, are older by a few months, are slightly more likely to be married, are slightly more educated, and are more likely to be employed when the child is 6 years old. However, the gross annual income of the parents seems to be similar between the two groups. In order to figure out how important these differences are for our educational attainment outcomes, in Panel D we compare the predicted educational outcomes between the two groups of children. The results, indicate that children who are prescribed AD tend

[14]We should note that our instrument is constructed using all 10–15 year old children who had a consultation with that specific specialist, not only those in our selected birth cohorts.

[15]In section 5 we test the sensitivity of our results to all the sample restrictions.

to be slightly positively selected, with higher predicted scores in both tests.

Panel E in the table lists some characteristics of the specialist providing the first consultation. On average, they tend to prescribe AD to about 17 percent of their patients, with slightly higher propensity to prescribe for the physicians treating children who end up receiving AD. Child psychiatrists consult around 115 children between 10 and 15 years old per year. Children who are prescribed an AD seem also to be more likely to have seen another specialist than the diagnosing physician before the age of 15, at a rate three times higher than children who do not fill an AD prescription. This underscores the potential endogeneity in provider choice and the importance of assigning children to the first provider who treats them. Panel F shows that around 44 percent of the children who eventually receive an AD prescription do so in the year of the first consultation, and around three quarters of these (0.325/0.435) are written by the specialist conducting this consultation. In total, children receive on average about 103 DDDs of AD between the ages of 10 and 15, that is, the equivalent of 3.5 months of treatment for an average adult.

Finally, Panel G describes the academic achievement of children in our sample. Roughly 77 percent of them take either test (the corresponding number in the population of children is who use psychiatric services is 71 percent), and they tend to perform worse than the average student in both Danish and math. As suggested before, children who fill an AD prescription tend to perform better than the rest of the children in the analysis sample, although part of this result could be due to selection because they are also less likely to take the two tests.

5 Results

5.1 Academic achievement

We start by describing the relationship between AD use and academic achievement. Panel A in Table 2 provides the estimated coefficients from simple regressions of 9th grade tests scores in Danish and math on the indicator for having taken AD at any point between the ages of 10 and 15. The specifications in columns 1 and 4 only control for child characteristics: immigrant status, gender, age at diagnosis, birth cohort, and municipality of residence at the time of diagnosis. Columns 2 and 5 add information on the mother in the year when the child is 6 years old: immigration and marital status, age, place in the distribution of total annual gross income. Finally, columns 3 and 6 add the corresponding variables for the fathers.

There are two main conclusions that can be drawn from the results in Panel A. First, adding parental controls significantly changes the estimated coefficients. In the case of Danish, controlling for parental characteristics reduces the correlation between using AD and test scores by about one third, while in the case of math it increases it by a factor of 5. Second, the specification including all the controls suggests that taking AD is associated with higher test scores in Danish by a statistically significant 10 percent of a standard deviation and lower in math by about 5 percent.

In order for an IV approach to provide useful estimates, the instrument needs to be relevant. In our case, this means that the specialist propensity to prescribe must have an effect on the probability of ever filling an AD prescription. We start by providing some visual evidence on this relationship in Figure 1. Figure 1(a) shows that there is a strong positive relationship between a specialist's propensity to prescribe AD to other children 10–15 years old and the probability of having a prescription filled at any point between the ages of 10–15. The same relationship can also be seen in terms of the total number of DDDs filled between the ages of 10 and 15, as shown in Figure 1(b).

Panel B in Table 2 investigates this relationship more formally. The estimates suggest that moving from the 25th (0.117) to the 75th (0.229) percentile in the distribution of physician propensity to prescribe increases the probability of ever using AD between the ages of 10 and 15 by 4.8–5.0 percentage points, or 13–14 percent as compared to the average probability. It is also important to note that the estimated coefficients are much more stable to the inclusion of covariates than the coefficients in Panel A, suggesting that our instrument is less likely to be correlated with any omitted variables. In addition, the F test of significance of the instrument produces a statistic that is well above 10, the rule-of-thumb threshold for a strong instrument.

We next examine the relationship between our instrument and the main outcomes, or the reduced form. As before, we start by presenting some visual evidence in Figure 2. Both in the case of Danish and in the case of math, the Figure suggests that higher propensity to prescribe AD to other children is associated with slightly higher test scores. This conclusion is supported by the estimated results in Panel C of Table 2, which indicate that moving from the 25th to the 75th percentile in the distribution of physician propensity to prescribe increases test scores by 0.019–0.026 (Danish) or 0.054–0.059 (math) percent of a standard deviation, with only the second set of coefficients being statistically significant. Just as in the case of the first stage, the estimates change little when additional controls are included, suggesting again that there is little correlation between the instrument and any relevant omitted variables.

Finally, the IV estimates of the effect of using AD on academic achievement are presented in Panel D of Table 2. In our one endogenous variable-one instrument framework, the IV coefficient is just the ratio of the reduced form estimate to the first stage estimate. It is not surprising, then, that we find that using AD leads to an insignificant increase in Danish test scores of approximately 0.4 of a standard deviation, and to a significant increase in math test scores of around 1.2 of a standard deviation. It is also not surprising that these results are relatively large and somewhat imprecise, but they are very stable to the the inclusion of controls, suggesting again that our IV approach is able to eliminate the effects of unobserved characteristics.

5.2 Instrument validity

We already showed in the previous section that our instrument satisfies the relevance condition. In the rest of this section we will investigate the remaining assumptions required for a valid instrument: the exogeneity assumption, the exclusion restriction, and the monotonicity assumption. Both of them are technically untestable, but we can bring some suggestive evidence on their plausibility.

The exogeneity assumption states that the instrument should be as good as randomly assigned. This means that the instrument should not be correlated with any unobserved characteristics of the individuals. In practice we can only check if the instrument is correlated with observed characteristics. If we do not find evidence of such a correlation, this makes us more confident that the instrument is also uncorrelated with the unobservables. We estimate regressions of all the covariates included in our main specification on the instrument and present the results in Table 3. Given that the "randomization" of children to specialists is done at the level of year of diagnosis and municipality of residence, all our regressions control for indicators for these two characteristics. Out of 20 estimates in the Table, only one is statistically significant at a 5 percent level and one more at a 10 percent level. Most of the estimates imply relatively small effects within the range of the instrument. For example, moving from the 25th to the 75th percentile in the distribution of propensity to prescribe AD is associated with an increase in the probability that the mother is married when the child is 6 years old of 2.3 percentage points, or 3.7 percent as compared to the mean. Similar to the approach in section 4, we gauge the importance of these correlations by regressing predicted educational outcomes on the instrument. The results in Panel E of Table 3 indicate again that the propensity to prescribe AD is not significantly related to the characteristics that matter for educational achievement. We interpret these findings to suggest that the instrument is likely uncorrelated with any relevant omitted variables.

The exclusion restriction holds that the instrument affects the outcomes of interest only through the endogenous variable. In our context, this implies, for example, that the propensity to prescribe AD should not convey any information about the general quality of care provided by the physician. In a future version we will provide some suggestive evidence on the plausibility of this assumption by checking if the educational outcomes of children who had a first contact with a child psychiatrist before the age of 10 (and who are thus more likely to suffer from ADHD and less likely to have been diagnosed with an emotional disorder) are correlated with our instrument.

The final requirement for a valid instrument is the monotonicity assumption, which requires the instrument to affect all individuals in the same way. In other words, all children should be more likely to be prescribed AD if they visit a specialist with higher propensity to prescribe AD. In the spirit of (Norris, 2019; Norris et al., 2019), in a future version we will check whether the first-stage coefficient of the instrument has a different sign in any subgroup as defined by several characteristics.

5.3 Robustness checks

In this section we check the sensitivity of our results to several choices made in the construction of the sample or of the instrument. We start by examining whether any of our sample selection criteria affect the baseline estimates. We report the estimates from IV models in Table 4 and the first stage and reduced form results in Appendix Table ??.

Our analysis sample includes children first treated by child psychiatrists who treat at least 10 children during the year of that first consultation. In Panel A of Table 4 we explore how the selection of specialists affects our results by requiring that they treat at least 20 or 50 children in the year of the first diagnosis. We find remarkably similar estimated effects, especially for math test scores, although the coefficients lose some statistical significance due to the smaller sample sizes. We then add to the sample children treated by general psychiatrists (who, again, have to treat at least 10 children 10–15 years old in that year). Although the estimated coefficients become smaller, we still find significant positive effects of AD treatment on math test scores.

In Panel B of the Table we turn to the role of the criteria used in selecting the children into the sample. In particular, we check whether the age limit we impose for the first psychiatric contact has any influence on the estimated effects of AD treatment. As we increase the minimum age for selection into the sample, we increase the likelihood that children are diagnosed with an emotional disorder but we also reduce the sample size and hence the power. The results indicate that AD use before the age of 16 is indeed related to higher test scores, particularly in math, with coefficients generally similar to our baseline estimates, although we do lose significance in the most restrictive sample (children at least 13 years old at diagnosis).

A body of literature in statistics discusses the advantages and benefits of different approaches to handling missing values of covariates (see Gelman and Hill, 2006, Chapter 25, for a review). In our main specification we employ the widely-used approach of imputing a value (in our case, the mode in the sample) to missing covariates and including an indicator for whether such an imputation was conducted for each variable. We report the results from two alternative approaches in Panel C of Table 4. First, we simply exclude observations with missing values for parental characteristics (the "complete-case" approach), which yields consistent estimates under the assumption that the values are missing at random. Second, we add to our main specification interactions between the indicators for missing values and the imputed variables (Jones, 1996). In both cases, we find that the effect of using AD is very similar to our baseline estimates, especially in the case of math test scores.

We next consider the possibility that our results are driven by outliers, i.e., by children who are treated by specialists whose propensity to prescribe AD lies in the tails of the distribution of propensity to prescribe. We apply to widely used techniques for dealing with outliers, namely excluding these observations and winsorizing the distribution (i.e., censoring the distribution at particular values in the tails). We examine the robustness of our results to excluding or winsorizing the 1 and 5 percent tails in Panel D of Table 4. As before, we find estimates that are very similar to our baseline effects of using AD, especially in the case of math test scores.

In Panel E of the Table we investigate alternative ways of constructing our instrument. Our main instrument is calculated as the propensity to prescribe AD to other children treated during the year of diagnosis, even if they were treated after the "focal child." Our results could then be biased if there is feedback from the treatment of the focal child to these other children. Although this scenario is unlikely to drive our results because of the generally long lag between treatment initiation and alleviation of symptoms (Trivedi et al., 2006), we report estimates from a specification using the physician propensity to prescribe AD in the calendar year prior to the first contact. The effects are more imprecisely estimated due to the smaller sample, but very close in magnitude to our baseline estimates. This confirms our conjecture that there is no within-year feedback that influences the propensity to prescribe AD of a specialist.

Our estimated effects represent a local average treatment effect (LATE) for the compliers, the children who receive AD because they are treated by a physician with a higher propensity to prescribe medication. The compliers cannot be identified in the data, but their characteristics can be described if the instrument is binary.[16] To this end, we first check if our results are similar when using a discrete version of our instrument, an indicator for whether the propensity to prescribe of the treating specialist is higher than the median propensity to prescribe in our sample. The estimated effects of using AD, shown in the bottom of Panel E of Table 4, are again similar to the baseline estimate in the case of math test scores.

Our results may be driven by the particular measure of treatment with AD used in the baseline specification. In Panel F of Table 4, we report estimates from specifications using two alternative measures of treatment with AD: the average number of DDDs filled per year between the year of diagnosis and the year the child turns 15, and an indicator for whether an AD prescription was filled in the year of the first contact with a child psychiatrist. We again find that treatment with AD has a statistically significant effect on math test scores. Receiving an AD prescription in the year of diagnosis leads to an increase in the math test score of 0.776 of a standard deviation, while an additional DDD per year increases it by 0.004 of a standard deviation. Among children in our sample who use AD, the 25th and the 75th percentile of the average yearly DDD are 32.67 and 182, respectively. Our estimates then suggest that moving from the 25th to the 75th percentile in terms of AD use increases math test scores by 0.616 of a standard deviation.

Finally, recall that students can be exempted from taking the standardized exit tests. In our IV framework we find that AD use leads to a decline in the probability of taking the exit tests of

[16]This method can also be applied in the case of multiple binary instruments, for example when a continuous-valued instrument is discretized. However, we choose to avoid this approach because of the issues related to multiple (weak) instruments (Bound et al., 1995; Stock et al., 2002).

27–29 percent. Although statistically insignificant, these results are large enough to warrant closer examination. If treatment with AD allows worse-performing students to be exempt from the test, or alternatively if it enables academically-better students to take the test, then our results would provide an upward-biased estimate of the true effect of using AD in our population of students. We examine this issue in several ways. First, we compare the predicted outcomes among non-test-takers who take and who do not take AD. The means listed in Table 5 show that the sample of non-test-takers is indeed negatively selected, although the pattern does not seem to differ between the groups of children who take and who do not take AD. Second, we predict the outcomes of non-test-takers using the test-takers in our sample and we re-estimate our baseline specification. The estimated effect of using AD is now 0.284 (s.e. 0.326) for Danish test scores and 0.814 (s.e. 0.372) for math test scores. Although lower than our baseline estimates, these findings still suggest that treatment with AD has significant benefits for academic achievement, especially with respect to math.

5.4 Academic achievement by gender

Previous research finds that girls are more likely than boys to be diagnosed with emotional disorders during adolescence (Costello et al., 2005; Thapar et al., 2012; Wesselhoeft et al., 2015; National Health Service, 2018). It would be interesting then to see if girls also benefit more from pharmaceutical treatment of their mental disorders. Unfortunately, our instrument is not strong enough in the sub-samples defined by gender to allow us to draw definitive conclusions. The results shown in Table 6 point again to larger effects on math test scores, particularly for girls, but they are very imprecisely estimated. Based on these, we cautiously conclude that girls seem to indeed benefit more than boys from AD.

6 Conclusions

In this paper we investigate the effects of AD use in adolescence on academic achievement, using the propensity to prescribe aD of the first specialist as instrument. our results point to large benefits from AD treatment, particularly for math tests. We also provide some suggestive evidence that girls benefit more than boys from AD treatment. These results are particularly important given the current debate on the effectiveness of the pharmaceutical treatment of depression and anxiety with AD.

References

Alexandre, Pierre K., and Michael T. French (2001), "Labor supply of poor residents in metropolitan Miami, Florida: The role of depression and the co-morbid effects of substance use," *Journal of*

Mental Health Policy and Economics 4(4), 161–173.

Amsterdam, Jay D. (1998), "Selective serotonin reuptake inhibitor efficacy in severe and melancholic depression," *Journal of Psychopharmacology* 12(4 suppl), S99–S111.

Andersen, John Sahl, Niels de Fine Olivarius, and Allan Krasnik (2011), "The Danish National Health Service Register," *Scandinavian Journal of Public Health* 39(7_suppl), 34–37.

Anderson, D. Mark, Resul Cesur, and Erdal Tekin (2015), "Youth depression and future criminal behavior," *Economic Inquiry* 53(1), 294–317.

Baadsgaard, Mikkel, and Jarl Quitzau (2011), "Danish registers on personal income and transfer payments," *Scandinavian Journal of Public Health* 39(7_suppl), 103–105.

Bauer, Jeanett, Niels Okkels, and Povl Munk-Jørgensen (2012), "State of psychiatry in Denmark," *International Review of Psychiatry* 24(4), 295–300.

Berndt, Ernst R., Stan N. Finkelstein, Paul E. Greenberg, Robert H. Howland, Alison Keith, A. John Rush, James Russell, and Martin B. Keller (1998), "Workplace performance effects from chronic depression and its treatment," *Journal of Health Economics* 17(5), 511–535.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson (2013), "Early life health interventions and academic achievement," *American Economic Review* 103(5), 1862–1891.

Bound, John, David A. Jaeger, and Regina M. Baker (1995), "Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak," *Journal of the American Statistical Association* 90(430), 443–450, articleType: primary_article / Full publication date: Jun., 1995 / Copyright © 1995 American Statistical Association.

Busch, Susan H., Ezra Golberstein, and Ellen Meara (2014), "The FDA and ABCs: Unintended consequences of antidepressant warnings on human capital," *Journal of Human Resources* 49(3), 540–571, volume 49, Number 3, Summer 2014.

Chatterji, Pinka, Margarita Alegria, and David Takeuchi (2011), "Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication," *Journal of Health Economics* 30(5), 858–868.

Costello, E. Jane, Helen L. Egger, and Adrian Angold (2005), "The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity," *Child and Adolescent Psychiatric Clinics of North America* 14(4), 631–648.

Currie, Janet, and Mark Stabile (2009), "Mental health in childhood and human capital," in Jonathan Gruber, ed., *The Problems of Disadvantaged Youth: An Economic Perspective*, 115 – 148, University of Chicago Press.

Dalsgaard, Søren, Helena Skyt Nielsen, and Marianne Simonsen (2014), "Consequences of ADHD medication use for children's outcomes," *Journal of Health Economics* 37, 137–151.

Doyle, Joseph J. (2007), "Child protection and child outcomes: Measuring the effects of foster care," *American Economic Review* 97(5), 1583–1610.

——— (2008), "Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care," *Journal of Political Economy* 116(4), 746–770.

Duggan, Mark (2005), "Do new prescription drugs pay for themselves?: The case of second-generation antipsychotics," *Journal of Health Economics* 24(1), 1–31.

Ettner, Susan L., Richard G. Frank, and Ronald C. Kessler (1997), "The impact of psychiatric disorders on labor market outcomes," *Industrial and Labor Relations Review* 51(1), 64–81.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth (2014), "The effects of poor neonatal health on children's cognitive development," *American Economic Review* 104(12), 3921–3955.

Fletcher, Jason (2013), "Adolescent depression and adult labor market outcomes," *Southern Economic Journal* 80(1), 26–49.

Fletcher, Jason M. (2010), "Adolescent depression and educational attainment: Results using sibling fixed effects," *Health Economics* 19(7), 855–871.

Fletcher, Jason M., and Steven F. Lehrer (2011), "Genetic lotteries within families," *Journal of Health Economics* 30(4), 647–659.

Galárraga, Omar, David S. Salkever, Judith A. Cook, and Stephen J. Gange (2010), "An instrumental variables evaluation of the effect of antidepressant use on employment among hiv-infected women using antiretroviral therapy in the united states: 1996–2004," *Health Economics* 19(2), 173–188.

Gelman, Andrew, and Jennifer Hill (2006), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge, New York: Cambridge University Press.

Green, Ben (2003), "Focus on paroxetine," *Current Medical Research and Opinion* 19(1), 13–21.

Jensen, Vibeke M., and Astrid W. Rasmussen (2011), "Danish education registers," *Scandinavian Journal of Public Health* 39(7_suppl), 91–94.

Jones, Michael P. (1996), "Indicator and stratification methods for missing explanatory variables in multiple linear regression," *Journal of the American Statistical Association* 91(433), 222–230.

Kessler, R. C., C. L. Foster, W. B. Saunders, and P. E. Stang (1995), "Social consequences of psychiatric disorders, I: Educational attainment," *American Journal of Psychiatry* 152(7), 1026–1032.

Kessler, Ronald C., Matthias Angermeyer, James C. Anthony, Ron De Graaf, Koen Demyttenaere, Isabelle Gasquet, Giovanni DE Girolamo, Semyon Gluzman, Oye Gureje, Josep Maria Haro, Norito Kawakami, Aimee Karam, Daphna Levinson, Maria Elena Medina Mora, Mark A. Oakley Browne, José Posada-Villa, Dan J. Stein, Cheuk Him Adley Tsang, Sergio Aguilar-Gaxiola, Jordi Alonso, Sing Lee, Steven Heeringa, Beth-Ellen Pennell, Patricia Berglund, Michael J. Gruber, Maria Petukhova, Somnath Chatterji, and T. Bedirhan Ustün (2007), "Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative," *World Psychiatry* 6(3), 168–176.

Kildemoes, Helle Wallach, Henrik Toft Sørensen, and Jesper Hallas (2011), "The Danish National Prescription Registry," *Scandinavian Journal of Public Health* 39(7_suppl), 38–41.

Ludwig, Jens, and Dave E. Marcotte (2005), "Anti-depressants, suicide, and drug regulation," *Journal of Policy Analysis and Management* 24(2), 249–272.

Ludwig, Jens, Dave E. Marcotte, and Karen Norberg (2009), "Anti-depressants and suicide," *Journal of Health Economics* 28(3), 659–676.

Lynge, Elsebeth, Jakob Lynge Sandegaard, and Matejka Rebolj (2011), "The Danish National Patient Register," *Scandinavian Journal of Public Health* 39(7_suppl), 30–33.

Maestas, Nicole, Kathleen J. Mullen, and Alexander 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.

Marcotte, Dave E., and Sara Markowitz (2011), "A cure for crime? Psycho-pharmaceuticals and crime trends," *Journal of Policy Analysis and Management* 30(1), 29–56.

Marcotte, Dave E., and Virginia Wilcox-Gök (2003), "Estimating earnings losses due to mental illness: A quantile regression approach," *Journal of Mental Health Policy and Economics* 6(3), 123–134.

Mors, Ole, Gurli P. Perto, and Preben Bo Mortensen (2011), "The Danish Psychiatric Central Research Register," *Scandinavian Journal of Public Health* 39(7_suppl), 54–57.

National Health Service (2018), "Mental health of children and young people in England, 2017," available online at https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017, last accessed on July 15, 2019.

Nemeroff, Charles B. (2007), "The burden of severe depression: A review of diagnostic challenges and treatment alternatives," *Journal of Psychiatric Research* 41(3), 189–206.

Norris, Samuel (2019), "Examiner inconsistency: Evidence from refugee appeals," Manuscript.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver (2019), "The effects of parental and sibling incarceration: Evidence from Ohio," Manuscript.

Pedersen, Carsten Bøcker (2011), "The Danish civil registration system," *Scandinavian Journal of Public Health* 39(7_suppl), 22–25.

Pedersen, Kjeld Møller, Terkel Christiansen, and Mickael Bech (2005), "The Danish health care system: Evolution—not revolution—in a decentralized system," *Health Economics* 14(S1), S41–S57.

Petersson, Flemming, Mikkel Baadsgaard, and Lau Caspar Thygesen (2011), "Danish registers on personal labour market affiliation," *Scandinavian Journal of Public Health* 39(7_suppl), 95–98.

Ryan, Neal D. (2003), "Medication treatment for depression in children and adolescents," *CNS Spectrums* 8(04), 283–287.

Schoenbaum, Michael, Jürgen Unützer, Daniel McCaffrey, Naihua Duan, Cathy Sherbourne, and Kenneth B Wells (2002), "The effects of primary care depression treatment on patients' clinical status and employment," *Health Services Research* 37(5), 1145–1158.

Skovlund, Charlotte Wessel, Lars Vedel Kessing, Lina Steinrud Mørch, and Øjvind Lidegaard (2017), "Increase in depression diagnoses and prescribed antidepressants among young girls. a national cohort study 2000–2013," *Nordic journal of psychiatry* 71(5), 378–385.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002), "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics* 20(4), 518–529.

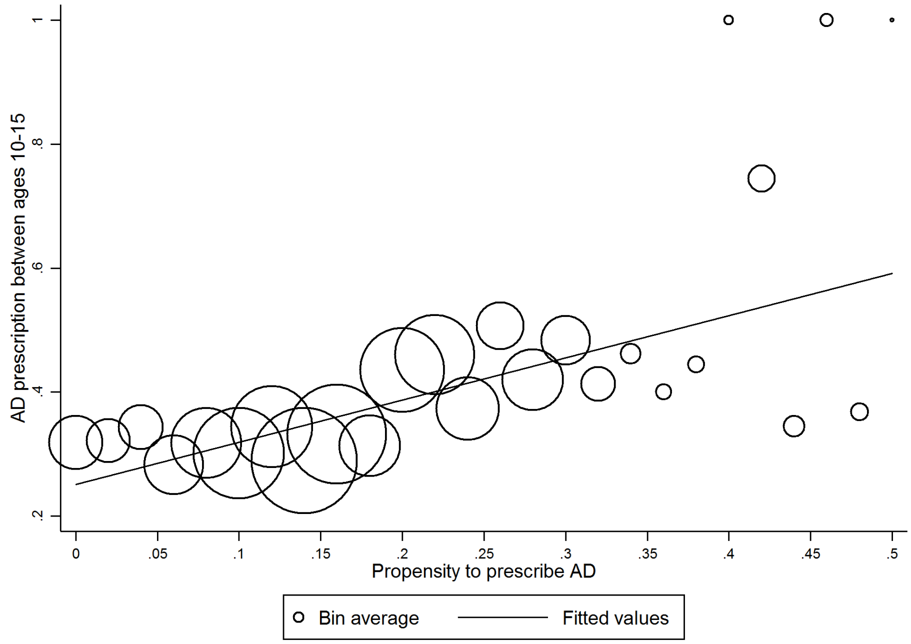
Thapar, Anita, Stephan Collishaw, Daniel S. Pine, and Ajay K. Thapar (2012), "Depression in adolescence," *Lancet* 379(9820), 1056–1067.

Trivedi, Madhukar H., Maurizio Fava, Stephen R. Wisniewski, Michael E. Thase, Frederick Quitkin, Diane Warden, Louise Ritz, Andrew A. Nierenberg, Barry D. Lebowitz, Melanie M. Biggs, James F. Luther, Kathy Shores-Wilson, A. John Rush, and STAR*D Study Team (2006), "Medication augmentation after the failure of SSRIs for depression," *New England Journal of Medicine* 354(12), 1243–1252.

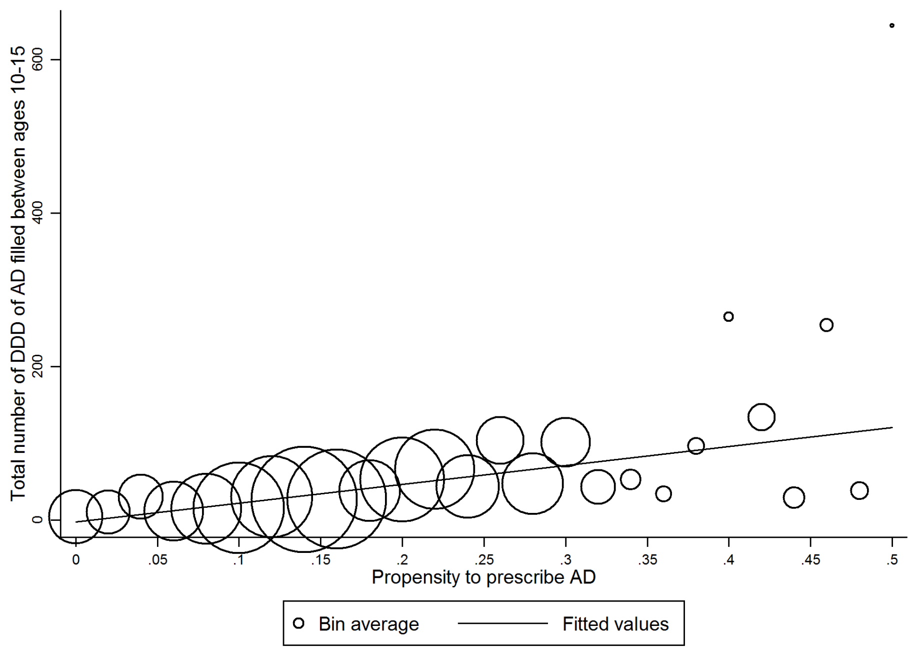
Wesselhoeft, R., C. B. Pedersen, P. B. Mortensen, O. Mors, and N. Bilenberg (2015), "Gender-age interaction in incidence rates of childhood emotional disorders," *Psychological Medicine* 45(4), 829–839.

WHO (2017), *Depression and Other Common Mental Disorders: Global Health Estimates*, Geneva: World Health Organization.

——— (2018), "Mental health in primary care: Illusion or inclusion?" WHO Technical Series on Primary Health Care WHO/HIS/SDS/2018.38.



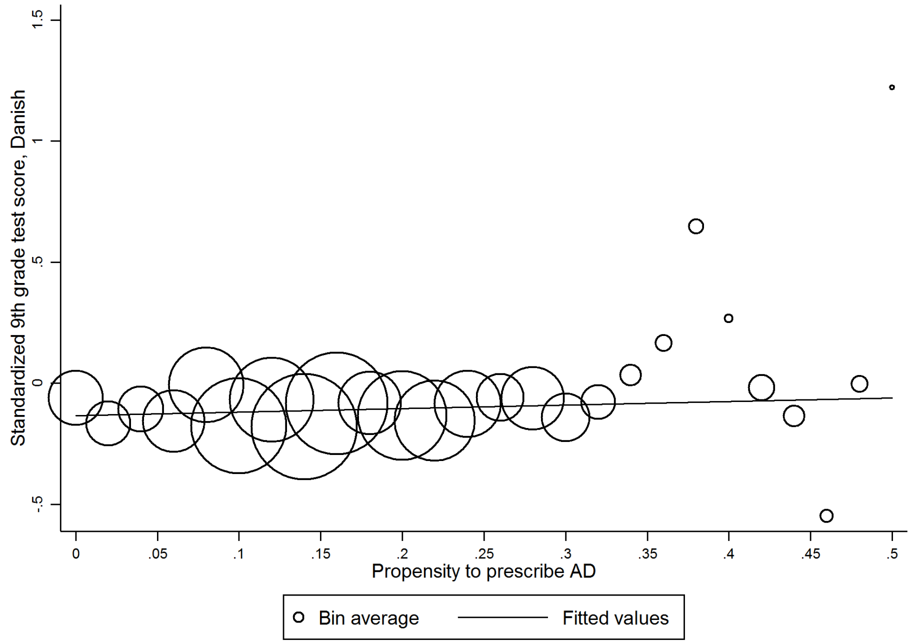
(a) Probability of ever filling a prescription between ages 10–15



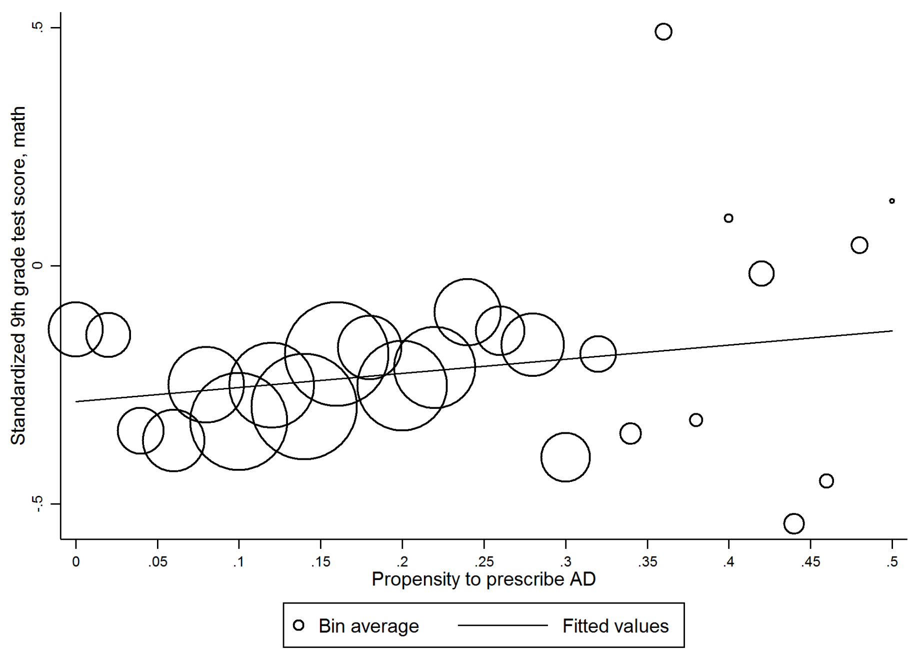
(b) Number of DDDs of AD filled between ages 10–15

Figure 1: Relationship between AD prescriptions and specialist propensity to prescribe

Notes: Each circle represents the average over a bin of width 0.02, and its size indicates the number of observations in the bin. The solid line plots the estimated simple relationship between the outcome and the physician propensity to prescribe.



(a) 9th grade test score, Danish



(b) 9th grade test score, math

Figure 2: Relationship between academic achievement and specialist propensity to prescribe

Notes: Each circle represents the average over a bin of width 0.02, and its size indicates the number of observations in the bin. The solid line plots the estimated simple relationship between the outcome and the physician propensity to prescribe.

Table 1: Sample characteristics

	All	Analysis sample			
		All	No AD	AD	p-value
	(1)	(2)	(3)	(4)	(5)
A. Child characteristics					
Age at diagnosis	13.362	13.033	12.904	13.260	0.000***
	(1.421)	(1.437)	(1.428)	(1.426)	
B. Mother's characteristics at child age 6					
Age	34.548	35.124	35.018	35.311	0.038**
	(5.040)	(5.028)	(5.091)	(4.911)	
Years of schooling	13.107	13.624	13.539	13.776	0.001***
	(2.400)	(2.485)	(2.487)	(2.473)	
Employed	0.718	0.768	0.750	0.798	0.000***
Missing values for any variable	0.048	0.041	0.044	0.034	0.069*
C. Father's characteristics at child age 6					
Age	37.278	37.588	37.414	37.894	0.004***
	(5.931)	(5.864)	(5.833)	(5.906)	
Years of schooling	13.452	13.888	13.845	13.965	0.107
	(2.578)	(2.619)	(2.596)	(2.657)	
Employed	0.853	0.878	0.870	0.892	0.017**
Missing values for any variable	0.088	0.079	0.089	0.063	0.000***

Table 1: Sample characteristics (cont.)

	All	Analysis sample			
		All	No AD	AD	p-value
	(1)	(2)	(3)	(4)	(5)
D. Predicted outcomes					
9th grade test score, Danish	1.514	1.554	1.503	1.643	0.000***
	(0.401)	(0.405)	(0.402)	(0.396)	
9th grade test score, math	2.205	2.296	2.284	2.317	0.003***
	(0.373)	(0.376)	(0.380)	(0.368)	
E. Specialist characteristics					
Physician propensity to prescribe AD in the year of diagnosis	—	0.173	0.165	0.188	0.000***
		(0.087)	(0.082)	(0.093)	
Physician propensity to prescribe AD in the year prior to diagnosis	—	0.173	0.167	0.182	0.000***
		(0.088)	(0.084)	(0.093)	
Number of 10–15 year old children treated in year of diagnosis	473.355	113.562	118.632	104.621	0.000***
	(514.924)	(65.866)	(69.043)	(58.812)	
Number of 10–15 year old children treated in year prior to diagnosis	440.139	107.107	111.921	98.424	0.000***
	(493.181)	(64.186)	(67.082)	(57.610)	
Indicator for changing doctors	0.173	0.162	0.089	0.290	0.000***
F. Filled AD prescriptions					
AD prescription between ages 10–15	0.332	0.362	0.000	1.000	—
AD prescription in year of diagnosis	—	—	—	0.435	—
AD prescription in year of diagnosis written by diagnosing specialist	—	—	—	0.325	—
Total number of DDD of AD filled between ages 10–15	—	—	—	103.253	—
				(181.454)	
G. Academic achievement					
Took 9th grade Danish test	0.711	0.776	0.782	0.767	0.211
Took 9th grade math test	0.710	0.766	0.774	0.752	0.068*
Age when test taken	16.862	16.824	16.839	16.796	0.105
	(0.808)	(0.795)	(0.770)	(0.837)	
9th grade test score, Danish	−0.171	−0.110	−0.215	0.078	0.000***
	(1.035)	(1.030)	(1.025)	(1.013)	
9th grade test score, math	−0.372	−0.237	−0.249	−0.215	0.294
	(1.020)	(1.011)	(1.033)	(0.971)	
Observations	38,412	5,373	3,429	1,944	

Notes: Column 1 reports statistics for the sample of children born between 1986–1999 who have their first contact with a hospital psychiatric department or a (child) psychiatrist between the ages of 10–15. Columns 2–5 report statistics for the analysis sample, i.e., the sample of children whose first contact is with a child psychiatrist who treated at least 10 children 10–15 years old during the year of that first consultation. Column 2 reports means for the entire analysis sample, column 3 for the subsample of children who do not fill an AD prescription between the ages of 10–15, and column 4 for the subsample of children who fill at least one AD prescription between the ages of 10–15. Column 5 reports the p-value for the test of equality between the means in columns 3 and 4. The predicted outcomes reported in panel D are calculated based on the entire population of 10–15 year-old children during our sample period. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Antidepressant use and academic achievement

	9th grade test score, Danish ($N = 4,172$)			9th grade test score, math ($N = 4,114$)		
	(1)	(2)	(3)	(4)	(5)	(6)
A. OLS						
Ever used AD	0.141***	0.104***	0.099***	−0.009	−0.050	−0.054
	(0.035)	(0.034)	(0.035)	(0.037)	(0.035)	(0.035)
Mean outcome	−0.215	−0.215	−0.215	−0.249	−0.249	−0.249
B. First stage						
Physician propensity to prescribe	0.443***	0.414***	0.415***	0.449***	0.426***	0.430***
	(0.097)	(0.097)	(0.098)	(0.099)	(0.099)	(0.101)
Mean outcome	0.357	0.357	0.357	0.355	0.355	0.355
First-stage F statistic	20.670	18.157	17.878	20.542	18.618	18.182
C. Reduced form						
Physician propensity to prescribe	0.231	0.178	0.170	0.530**	0.485**	0.500**
	(0.198)	(0.188)	(0.191)	(0.209)	(0.202)	(0.202)
Mean outcome	−0.110	−0.110	−0.110	−0.237	−0.237	−0.237
D. IV						
Ever used AD	0.522	0.429	0.409	1.181**	1.137**	1.162**
	(0.469)	(0.474)	(0.474)	(0.550)	(0.568)	(0.570)
Mean outcome	−0.215	−0.215	−0.215	−0.249	−0.249	−0.249
Child characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Mother's characteristics	No	Yes	Yes	No	Yes	Yes
Father's characteristics	No	No	Yes	No	No	Yes

Notes: Sample of children whose first contact is with a child psychiatrist who treated at least 10 children 10–15 years old during the year of that first consultation. Each cell reports the estimated coefficient from a separate regression with outcome variable indicated in the column heading, dependent variable indicated in the row heading, specification indicated in the panel heading, and list of control variables indicated in the footer. Mean outcomes are calculated for children not using AD (Panels A and D) or for all children in the sample (Panels B and C). Child characteristics include indicators for whether the child is an immigrant (first- or second-generation), boy, for the age at first contact with a specialist, for the year of birth, and for the municipality of residence at the time of diagnosis. Mother's and father's characteristics include the number of years of schooling and indicators for whether they are an immigrant (first- or second-generation), married, for their age, and for the total annual gross income decile in the full population, all measured in the year when the child is 6 years old. Information on a small number of parents is not available, in which case the corresponding variables are set equal to the mode in the analysis sample and indicators for missingness are included in the regression. Robust standard errors clustered at the specialist-year of diagnosis level are reported in brackets. First-stage F statistic refers to the F-test of significance of the instrument in the specification in Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Instrument balance

	(1)	(2)	(3)	(4)	(5)	(6)
			A. Child characteristics			
	Boy	Immigrant	Age at diagnosis	Number of siblings	Number of older siblings	Number of younger siblings
Physician propensity to prescribe	−0.133	−0.038	0.174	−0.086	−0.129	0.043
	(0.099)	(0.033)	(0.252)	(0.180)	(0.161)	(0.159)
Observations	5,373	5,373	5,373	5,370	5,370	5,370
Mean outcome	0.501	0.054	13.358	1.558	0.770	0.789
			B. Mother's characteristics at child age 6			
	Age	Immigrant	Married	Years of schooling	Employed	Gross annual income
Physician propensity to prescribe	0.707	0.024	0.204**	0.018	0.136*	−1.206
	(0.991)	(0.046)	(0.094)	(0.460)	(0.081)	(27.235)
Observations	5,373	5,373	5,373	5,373	5,373	5,286
Mean outcome	34.547	0.060	0.625	13.106	0.718	285.113
			C. Father's characteristics at child age 6			
	Age	Immigrant	Married	Years of schooling	Employed	Gross annual income
Physician propensity to prescribe	1.297	−0.026	0.108	−0.242	0.074	2.890
	(1.187)	(0.047)	(0.091)	(0.488)	(0.060)	(49.969)
Observations	5,373	5,373	5,373	5,373	5,373	5,165
Mean outcome	37.276	0.071	0.651	13.451	0.853	374.211
	D. Missing parental characteristics		**E. Predicted academic outcomes**			
	Mother	Father	9th grade test score, Danish	9th grade test score, math	Took test, Danish	Took test, math
Physician propensity to prescribe	−0.021	−0.049	0.081	0.034	0.002	−0.001
	(0.039)	(0.053)	(0.078)	(0.072)	(0.006)	(0.006)
Observations	5,373	5,373	4,866	4,866	3,926	3,924
Mean outcome	0.048	0.088	1.513	2.205	0.983	0.980

Notes: Sample of children whose first contact is with a child psychiatrist who treated at least 10 children 10–15 years old during the year of that first consultation. Each cell reports the estimated coefficient from a separate OLS regression with outcome variable indicated in the column heading and independent variables: physician propensity to prescribe and indicators for year of diagnosis and municipality of residence. Mean outcomes are calculated for the entire analysis sample. Robust standard errors are reported in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Robustness checks

	9th grade test scores	
	Danish	Math
	(1)	(2)
A. Selection of specialists		
Number of children treated in year of diagnosis ≥ 20	0.208	1.249*
	(0.507)	(0.640)
Observations	4,121	4,063
Mean outcome	−0.214	−0.251
First-stage F statistic	14.715	15.317
Number of children treated in year of diagnosis ≥ 50	0.330	1.183*
	(0.492)	(0.612)
Observations	3,710	3,663
Mean outcome	−0.224	−0.261
First-stage F statistic	18.079	18.987
Add children first treated by general psychiatrists	0.159	0.732*
	(0.355)	(0.414)
Observations	4,302	4,241
Mean outcome	−0.223	−0.260
First-stage F statistic	25.708	25.696
B. Selection of children		
First psychiatric consultation at or after age 12	0.287	1.744**
	(0.593)	(0.833)
Observations	3,342	3,281
Mean outcome	−0.229	−0.276
First-stage F statistic	12.141	11.020
First psychiatric consultation at or after age 13	0.373	1.114
	(0.623)	(0.748)
Observations	2,507	2,448
Mean outcome	−0.209	−0.278
First-stage F statistic	9.270	9.781
C. Handling of missing parental characteristics		
Exclude children with missing parental characteristics	0.431	1.259**
	(0.554)	(0.633)
Observations	3,823	3,776
Mean outcome	−0.205	−0.235
First-stage F statistic	13.979	14.155
Add interactions between indicators for missingness	0.344	1.075*
and the corresponding variables	(0.471)	(0.572)
Observations	4,172	4,114
Mean outcome	−0.215	−0.249
First-stage F statistic	18.321	17.843

Table 4: Robustness checks (cont.)

	9th grade test scores	
	Danish	Math
	(1)	(2)
D. Handling of outliers in the distribution of physician propensity to prescribe		
Exclude 1% tails	0.311	1.270*
	(0.511)	(0.646)
Observations	4,047	3,993
Mean outcome	−0.217	−0.253
First-stage F statistic	14.270	13.931
Exclude 5% tails	−0.034	1.324*
	(0.481)	(0.678)
Observations	3,791	3,746
Mean outcome	−0.222	−0.261
First-stage F statistic	15.339	12.621
Winsorize 1% tails	0.378	1.169**
	(0.480)	(0.582)
Observations	4,172	4,114
Mean outcome	−0.215	−0.249
First-stage F statistic	17.211	17.496
Winsorize 5% tails	0.341	1.233**
	(0.492)	(0.623)
Observations	4,172	4,114
Mean outcome	−0.215	−0.249
First-stage F statistic	16.144	15.699
E. Alternative instruments		
Propensity to prescribe in the year prior to diagnosis	0.404	1.059
	(0.616)	(0.909)
Observations	3,872	3,823
Mean outcome	−0.215	−0.249
First-stage F statistic	5.657	5.242
Physician propensity to prescribe above median	−0.033	1.385**
	(0.480)	(0.640)
Observations	4,172	4,114
Mean outcome	−0.215	−0.249
First-stage F statistic	18.887	15.947
F. Alternative measures of AD use		
Average yearly number of DDDs until age 15	0.001	0.004**
	(0.002)	(0.002)
Observations	4,172	4,114
Mean outcome	−0.161	−0.260
First-stage F statistic	24.824	25.518
AD prescription in year of first consultation	0.267	0.776**
	(0.301)	(0.332)
Observations	4,172	4,114
Mean outcome	−0.161	−0.260
First-stage F statistic	90.472	92.704

Notes: See notes to Table 2 for a description of the sample and control variables used.

Table 5: Predicted academic achievement, test takers and non-test-takers

	Test-takers			Non-test-takers		
No AD	AD	p-value	No AD	AD	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted 9th grade test score, Danish	1.556	1.692	0.000***	1.301	1.474	0.000***
	(0.386)	(0.382)		(0.396)	(0.398)	
Observations	2,429	1,392	3,821	643	402	1,045
Predicted 9th grade test score, math	2.331	2.361	0.015**	2.112	2.180	0.004***
	(0.362)	(0.352)		(0.396)	(0.383)	
Observations	2,411	1,364	3,775	661	430	1,091

Notes: See notes to Table 1 for a description of the sample and variables used. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Heterogeneity in the effects of AD use

	9th grade test scores	
	Danish	Math
	(1)	(2)
A. Child gender		
Boy	0.401	0.932
	(0.574)	(0.626)
Mean outcome	−0.495	−0.275
Observations	2,162	2,156
First-stage F statistic	13.536	11.981
Girl	0.375	1.783
	(0.822)	(1.081)
Mean outcome	0.202	−0.210
Observations	1,914	1,860
First-stage F statistic	5.379	6.208

Notes: See notes to Table 2 for a description of the sample and control variables used.