Juvenile Crime and the Medical Treatment of Attention Deficit Hyperactivity Disorder*

Anne Sofie Tegner Anker (asa@rff.dk)
ROCKWOOL Foundation

Peter Fallesen (peter.fallesen@sofi.su.se)
ROCKWOOL Foundation
Stockholm University
European University Institute

*Funding gratefully acknowledged from the ROCKWOOL Foundation (grant #1082) and the Swedish Research Council for Health, Working Life and Welfare (#2016-07099). Authorship is not shared equally between the authors—the first author did the lion's share of the work. We thank Lars H. Andersen, Kristin Turney, Paul Nieuwberta, Lasse Liebst, and Anders Holm for comments on earlier versions. Part of this work was carried out while Anne Sofie Tegner Anker was a graduate student at Department of Sociology, University of Copenhagen, Denmark.

ABSTRACT

Children with Attention Deficit Hyperactivity Disorder (ADHD) have on average worse socioeconomic outcomes in adolescence and early adulthood, as well as higher risk of engaging in criminal activities and having criminal justice contacts. However, little attention has been directed at understanding whether medical treatment of ADHD symptoms may lessen a criminogenic effect of ADHD. This study examines whether consistent medical treatment of ADHD decreases the risk of crime. We utilize the fact that the medication does not work for a subgroup of the treated, and that this medical non-response occurs in a plausibly random manner. Danish administrative data on purchased medication allow us to identify a group who most likely discontinue treatment due to medical non-response. We find that treatment discontinuation compared to following a consistent treatment pattern results in a 21% and 36% relative increase in the risk of being charged or convicted by age 18, respectively, after controlling for a wide range of covariates. The findings stand up to a series of sensitive tests, negative control analyses, and bounding exercises. In total, we find a crime-reducing effect of treatment for those who qualify for and respond medically to ADHD treatment. Medical treatment not only addresses core ADHD symptoms, but also can play a substantial role in mitigating the increased risk of criminal offending associated with ADHD, thereby lessening social inequality in individual life chances caused by the disorder.

Keywords: ADHD; administrative data; crime; medical treatment

Word count: 9909

INTRODUCTION

Adolescents and young adults with Attention-Deficit Hyperactivity Disorder (ADHD) are at substantial higher risk of adverse life outcomes, including contacts with the criminal justice system (e.g., Young et al. 2015; Mohr-Jensen et al. 2019; Dalsgaard et al. 2013). Core symptoms of ADHD include delay aversion, poorer behavioral inhibition, and failure to consider the consequences of actions, and individuals who present with ADHD more often engage in risky or poorly considered behavior (Schilling, Walsh, and Yun 2011; Dalsgaard et al. 2015). These defining ADHD symptoms overlap considerably with low self-control, one of the most acknowledged explanations of criminal behavior (Wright and Beaver 2005; Schilling, Walsh, and Yun 2011). If ADHD is an independent, criminogenic risk factor, correct treatment of the disorder could have crime-reducing effects. Medical treatment with both psychotropic and other types of medication is highly effective in reducing core symptoms (e.g., Vaughan, March, and Kratochvil 2012; Faraone et al. 2021), which might in turn reduce the risk of committing crime.

This study estimates the effect of medical treatment of ADHD treatment on crime by utilizing plausibly random variation in people's physical response to medical treatment. 25-35% of those who commence medical treatment of ADHD does not respond well to treatment (i.e. are "medical non-responders") either because of side effects or because the medication fails to reduce core symptoms (Gray & Kagan, 2000; Johnston, Coghill, Matthews, & Steele, 2015). The non-response appears random—no clinical, demographic or socio-economic factors consistently predicts it (Kim, Sharma, and Ryan 2015; Gray and Kagan 2000; Keilow, Holm, and Fallesen 2018). We utilize the plausibly random variation in (non-)response to medical treatment using individual-level longitudinal data on all purchases of ADHD medication in Denmark from 1995-2013 linked to demographic registries and criminal records. The data allow us to identify distinct patterns of treatment within the group of

children receiving medical ADHD treatment before the age of 15 (N=7390). Measuring crime as charges and convictions by the age of 18, we find that compared to children with a consistent treatment pattern, a patterns of discontinued treatment results in a 21% and 36% relative increase in the risk of being charged or convicted, respectively, after controlling for a wide range of covariates. The finding withstands a series of sensitivity, negative control (Arnold et al. 2016), and bounding analyses (Oster 2019), bolstering a causal interpretation of the findings. Thereby, we provide a plausible estimate of the average treatment effect of ADHD medication on juvenile crime among those diagnosed with ADHD and prescribed medication, a key margin not previously studied within the literature.

The results are in line with existing studies examining the effect of ADHD medication on crime at other margins (Dalsgaard, Nielsen, and Simonsen 2014; Lichtenstein et al. 2012; Mohr-Jensen et al. 2019) thus adding to the burden of proof that ADHD medication reduces crime for some populations. As institutional features may be important for how the link between ADHD medication and criminal behavior unfolds, results are likely to generalize to other contexts with treatment regimens and diagnostic manuals as found in Denmark but may generalize less well to contexts with radical different approaches to prescription of medical treatment for ADHD. The study contributes to the ongoing discussion on whether medical treatment of behavioral disorders should be seen as a medicalization of deviant behavior, or instead as a way to offset adverse life consequence caused at least partly by biological processes not fully within the individual's control.

BACKGROUND

The Link Between ADHD and Crime

Attention-Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder, defined by age-inappropriate inattention, hyperactivity and impulsivity with symptoms occurring

before the age of 12. Roughly 5-7% of children worldwide meet the diagnostic criteria (Polanczyk et al. 2007; Roberts, Milich, and Barkley 2015; Sayal et al. 2018). For adults, the prevalence of ADHD is reported to be 3-5% (Roberts, Milich, and Barkley 2015).¹ Yet, meta-analysis found the prevalence of ADHD in incarcerated populations to range from 25-43% depending on the mode of diagnosis (Young et al. 2015). Individuals diagnosed with ADHD have also been found to have a much higher rates of various forms of criminal justice contact (e.g., Pratt et al. 2002; Mohr-Jensen et al. 2019; Dalsgaard et al. 2013).

ADHD and self-control

There is substantial overlap between symptoms of ADHD and low self-control, one of the most prominent and acknowledged theoretical and empirical explanations of criminal behavior (Unnever, Cullen, and Pratt 2003; Pratt et al. 2002; Schilling, Walsh, and Yun 2011; Wright and Beaver 2005). Individuals with low self-control are “impulsive, insensitive, physical (as opposed to mental), risk-taking, short-sighted, and non-verbal” (Gottfredson and Hirschi 1990, p.90) and low self-control is described as the “tendency to pursue short-term immediate pleasure” (Gottfredson and Hirschi 1990 p.93). This characterization of low self-control bears a remarkable similarity to the description of ADHD symptoms (Roberts, Milich, and Barkley 2015), and enough so that some researchers have hypothesized that ADHD may be a cause of low self-control (Wright and Beaver 2005; Unnever, Cullen, and Pratt 2003; Schilling, Walsh, and Yun 2011).

There is, however, little overlap in the proposed etiologies of ADHD and low self-control. Gottfredson and Hirschi claim ineffective parental management (i.e., the inability to monitor, recognize, and punish deviant behavior) to be the single most important cause of low self-control. In contrast, ADHD has been shown to be highly hereditary with genetic factors explaining roughly 70-80% of the variation in symptoms (Faraone and Larsson 2019; Barkely 2015) leaving little room for environmental explanations for the occurrence of

ADHD.² Given the close relationship between ADHD and low self-control, the central role of genetic factors in the etiology of ADHD challenges the claim that low self-control is solely or primarily caused by inadequate parental management (Unnever, Cullen, and Pratt 2003). .

While the current study does not attempt to disentangle the relationship between ADHD and self-control, the resemblance in descriptions of characterizing behavior and the close empirical link between the two phenomena highlight ADHD's relevance to the study of crime.

Direct and indirect empirical mechanisms linking ADHD and crime

ADHD may increase a person's propensity to engage in criminal activities both through direct and indirect channels. Directly because the hallmark symptoms of ADHD – particularly hyperactivity and impulsivity – lead to risky and sensation-seeking behavior with little regard for rules or consequences. People with ADHD do not find events and situations as stimulating as their peers, which lead them to seek out pleasure and stimulus through risky situations and activities – some of which might be illegal (Schilling, Walsh, and Yun 2011). The propensity to engage in risky behavior is also reflected in an excess mortality among people with ADHD driven mainly by accidental deaths (Dalsgaard et al. 2015). Additionally, people with ADHD are less able to suppress pleasure and stimulation impulses (Schilling, Walsh, and Yun 2011), and show delayed brain maturation particularly in the prefrontal cortex, which plays a key part in complex cognitive processes, moral judgements, and in conforming to situational social expectations (Barkely 2015; Schilling, Walsh, and Yun 2011). Thus, part of the direct effect might be channeled through lower levels of self-control (Unnever, Cullen, and Pratt 2003).

ADHD may also indirectly influence the propensity to engage in crime, because the disorder elicits negative responses from the environment, which may act to (further)

marginalize people or have criminogenic influences in and of themselves. In school, children with (untreated) ADHD generally have worse academic and social outcomes than classmates (Keilow, Holm, and Fallesen 2018; McQuade and Hoza 2015). In the home, (untreated) ADHD increase risk parental divorce (Schermerhorn et al. 2012) and of being placed in foster care (Fallesen and Wildeman 2015). Children with ADHD can be more difficult to parent, which may elicit negative and harsher parenting strategies conducive to anti-social behavior (Schilling et al., 2011). These environmental responses may be significant drivers of potential criminogenic effects of ADHD.

Medical Treatment of ADHD

Although most often considered a chronic or long-term disorder (Sayal et al. 2018), ADHD symptoms are treatable through medication and therapy. The most prescribed medications for ADHD are psychotropic stimulants and atomoxetine. Both drug-types have been shown to reduce core symptoms of ADHD by increasing/facilitating dopamine and norepinephrine activity in the prefrontal cortex, which leads to increased attention and relieves feelings of boredom (Schilling, Walsh, and Yun 2011; Vaughan, March, and Kratochvil 2012). The alleviation of symptoms may translate to lower risks of crime directly by reducing the risky and sensation-seeking behavior conducive to crime, and indirectly by minimizing collateral consequences of ADHD in the school or home environment (Fallesen and Wildeman 2015; Keilow, Holm, and Fallesen 2018; Schermerhorn et al. 2012).

Prior studies of the effects of ADHD medication on crime

Existing studies on the effects of ADHD medication on crime point in the direction of a crime reducing effect of medication. A Danish study used hospital level propensity for prescribing ADHD medication as an exogenous shock to treatment probability and found significant reductions in police interactions before the age of 15, which is the age of criminal

responsibility in Denmark (Dalsgaard, Nielsen, and Simonsen 2014). Two studies from Denmark and Sweden compared the crime rates in periods with and without treatment for people receiving medical treatment and found substantial (roughly 30%) reductions in conviction risk when medicated compared to when not medicated (Lichtenstein et al. 2012; Mohr-Jensen et al. 2019). While both strategies move our knowledge of the potential of ADHD medication to reduce criminal behavior significantly forward, they have limitations. Relying on hospital variation is sensitive to sorting into regions or hospitals. Starting medical treatment and abstaining from crime may both be driven by other life stabilizing events. Furthermore, prior studies have identified effects for specific subgroups of treated children – the child at the margin of entering treatment (who enter treatment before age 10, which is relatively early) and for people who circle in and out of medication. Thus, examining this topic through a third lens—the average treatment effect on the treated—serves to triangulate and bolster our knowledge of effects of ADHD medication on crime.

Medical non-response to ADHD treatment, an unexplored margin

Despite the overall effectiveness of ADHD medication, findings from the medical literature shows that stimulants are only effective for 65-75% of the treated population (Johnston et al., 2015). A larger share of treated (80-90%) may respond if several drugs are tried, but for a considerable share treatment either is ineffective in reducing symptoms or causes severe side effects (Vaughan, March, and Kratochvil 2012). Studies have found no consistent predictors of medical non-response to ADHD medication (Pelham and Smith 2000; Gray and Kagan 2000; Keilow, Holm, and Fallesen 2018; Kim, Sharma, and Ryan 2015). Demographic characteristics like age, sex, and race do not consistently predict treatment response (Chazan et al., 2011; Gray & Kagan, 2000; Johnston et al., 2015; Kim et al., 2015) and nor do IQ, socio-economic characteristics, or measures of family functioning (Chazan et al., 2011; Gray & Kagan, 2000; Johnston et al., 2015). For initial symptom severity findings are mixed, as

both higher and lower symptom severity have been associated with worse responses to treatment (Vallejo-Valdivielso et al. 2020; de Sonnevile, Njokiktjien, and Bos 1994; Gray and Kagan 2000). In sum, the general consensus is that there are no consistent demographic, socio-economic or clinical predictors of medical non-response to ADHD medication (Gray and Kagan 2000; Pelham and Smith 2000; Keilow, Holm, and Fallesen 2018; Kim, Sharma, and Ryan 2015).

Previous work has argued that response to medical treatment is as good as random and used this to estimate effects of medical ADHD treatment on academic achievements in Denmark (Keilow, Holm, and Fallesen 2018). We use a similar research design to study the effects of ADHD medication on crime. By exploiting the longitudinal pattern of ADHD medication purchases, estimates from the current study are likely to reflect the effect of taking medication rather than being diagnosed and entering treatment (which may be consequential in and of itself). Rather than identifying effects on the marginal child entering treatment, the current study identifies the average treatment effect on the treated (ATT), which is of theoretical importance since it informs us of the *average* potential of medication to mitigate criminogenic effects of ADHD among those with the disorder.

ADHD diagnosis and treatment in Denmark

In Denmark, 3.7% of cohorts born between 1990-1999 are diagnosed with ADHD before the age of 18 and 3% have entered treatment (Keilow, Wu, and Obel 2020). However, these rates reflect a stark increase in both treatment and diagnosis since the early 2000s (Fallesen and Wildeman 2015), which mirrors a global tendency (Sayal et al. 2018). The prevalence of diagnosis and treatment are in line with what is estimated in several comparable countries, but higher than the U.K. and much lower than the U.S. (Sayal et al. 2018).

While the features of the pathways to ADHD diagnosis and treatment in Denmark mirrors many European countries, it diverges considerably from how ADHD is diagnosed and treated in other developed countries, including the U.S. First, in Denmark, ADHD is diagnosed using the ICD 10-diagnostic criteria, which is more restrictive than the DSM-6 diagnostic criteria employed in the U.S. Second, in Denmark, diagnostic decisions are made by specialist child psychiatrists rather than by general practitioners (Perrild 2020). Third, national recommendations suggest that first line of treatment for milder cases should be non-pharmacological interventions (i.e., cognitive programs, parent-training programs, or social skills programs). Pharmacological treatment is therefore mostly initiated for more severe cases of ADHD (Danish Health Authority 2014) in contrast to the U.S. where the first line of treatment is pharmacological (Sayal et al. 2018). Fourth, the Danish socialized medical system reimburses a substantial part of the cost of medication (and gives full reimbursement for low-income groups).

The Danish approach to diagnosis and treatment of ADHD have implications for the margin of treatment. The average (or marginal) child treated with ADHD medication will have more severe symptoms in Denmark than in the U.S. (but is likely similar to other European countries with similar systems and prevalence of diagnosis and treatment). Additionally, the risk of overmedication and strategic use of ADHD medication as a performance-enhancing drug is generally quite low. Studies from the U.S. and Iceland have found disproportionate medication of children who are young for their grade (Zoëga, Valdimarsdóttir, and Hernández-Díaz 2012; Elder 2010) and that ADHD medication use drops considerably during school holidays (King, Jennings, and Fletcher 2014). These patterns are not mirrored in Denmark (Dalsgaard et al. 2014). Thereby, Denmark provides a context for studying the average effect of medication on crime among those with severe enough ADHD symptoms to medically warrant treatment.

DATA

For this study, we use administrative data available for Denmark. The administrative data encompasses the entire population of residents linked across of a wide range of registers with information on health care services use, criminal justice contacts, demographic characteristics, educational attainment, income, and welfare services. All prescription medication purchases are linkable through the personal identifier. The data are housed and maintained by Statistics Denmark, who provides access to a deidentified version for research purposes.

Analytical Sample

We use birth cohorts from 1986-1996, for whom we observe criminal charge and conviction histories from age of criminal responsibility at 15 to age of legal adulthood at 18. From the Danish Prescription Database, we identify all purchases of medications used to treat ADHD [dexamphetamine (N06BA02), methylphenidate (N06BA04, e.g., Ritalin), atomoxetine (N06BA09) and lisdexamphetamine (N06BA12)] for the period 1995-2013. Among the selected cohorts, 3.3% fill prescriptions for these medications during this period, and 1.2 % do so before the age of 15. We limit the sample to children who started medical treatment no later than three months before age 15 and purchased medication at least three times before age 15 to increase the likelihood that stopping treatment is due to medical non-response rather than lower commitment to treatment or medical treatment aversion. The sample restrictions result in N=7390 (see Table A-1). Everyone in the analytical sample has been diagnosed, commenced treatment, and shown initial commitment to the treatment.

ADHD Medication

Following the categorization described by Keilow and colleagues (2018), we identify three patterns of medical treatment in the data. 1) *Discontinued treatment pattern* (DTP), who

purchased medication for a period of maximum 90 days within the observation window (prior to age 15). The 90-day cut-off allows time for an initial period of medication adjustment and dosage titration, based on which a decision to terminate treatment due to medical non-response might follow. Recommendations from the British National Institute for Health and Care Excellence, which is referenced by the Danish Health Authorities, sets the trial period of an ADHD medication to 6 weeks (NICE 2018), and a 90 day period thus allows for trials of two separate medications. 2) *Consistent treatment pattern* (CTP), who purchased medication at least every 90 days. Since there is a non-negligible risk that ADHD medication may be misused/abused, prescriptions and purchases are tightly monitored by health authorities (Region Midtjylland 2016). This limits the possibility of “stocking up” on medication, and thus requires medication purchases on a regular basis to sustain a consistent use of medication. 3) *Ambiguous treatment pattern* (ATP) captures the remaining individuals. This group is characterized by dropping in and out of treatment and having longer spells between purchases.

Figure 1 shows the longitudinal medication purchases of those treated before the age of 15—overall and by defined treatment pattern. By age 15, 17% had a consistent treatment pattern, 4 % had discontinued treatment pattern and the remaining majority (79 %) had an ambiguous treatment pattern. The share purchasing medication in each month hovers around 40% except from the first few months where the share is around 50-60% and the number of days between prescriptions range between 50-70. Figure 1(b) and (d) illustrate how these measures differ across the defined treatment patterns. The CTP group purchases medication more frequently compared to the ATP group (every 30 days on average compared to roughly 70). The DPT group only purchases medication within three months of the first purchase by definition and initially they purchase medication with the same time intervals as those with CTP.

[Figure 1 about here]

Crime Measures

We measure crime as formal charges and convictions reported by the National Danish Police occurring from ages 15 through 17 as both binary indicators and count variables. Convictions are, however, only measured until age 18 for birth cohorts 1986-1994 due to data restrictions, which results in N=4332 for these measures. Since we observe the charge and conviction measures at rather young ages, we do not make any restrictions on the type of crime and include both violations of the Penal Code and less serious violations of Special Acts (including minor drug offenses and traffic offenses). By relying on formal charges/convictions we naturally do not observe all criminal behavior, but only behavior that results in an institutional response. This is a limitation we address in the discussion.

Covariates

From the administrative data we draw an extensive set of control variables (all of which are included as covariates in the analysis):

Birth and demographic characteristics. For birth outcomes, we include indicators of initial health: low birth weight (=1 if <2500 g), gestational age (premature = 1 if <36 weeks), and 5-minute Apgar score to measure health immediately after birth (low Apgar=1 if <7). Birth outcomes correlate with ADHD prevalence (Barkely 2015) and crime, so may act as confounders. Besides birth outcomes, we also include gender, immigrant background, birth cohort, and residential region at age 6, which are all dimensions in which both crime and the prevalence and treatment of ADHD are known to differ. We also include a measure of family structure (lives with both parents, with a single parent, with a parent and a stepparent, or other) at age 6, as an indicator stability in the family.

Treatment characteristics. We include dummies for whether the child starts treatment at an age that was relatively young for one's birth cohort (≤ 25 percentile) or relatively old for one's birth cohort (≥ 75 percentile). We also include age- and year-adjusted measures for Defined Daily Dosage's (DDD) of the first and third medication purchase as dummies for purchasing medication of a relatively high dosage (≥ 75 percentile) and indicators for whether more than one prescription was filled at the first and third medication purchase. Age at treatment and DDD capture severity, since earlier treatment could reflect more pronounced and more easily diagnosed ADHD. The measures are scaled according to age, cohort, and year to adjust for changes in medication practices.

Other psychoactive medication. To control for comorbidities such as autism spectrum disorder and depression, we include a dummy for purchasing psychoactive medication for other psychiatric disorders by the age of 15.

Parental characteristics. We include measures of parental characteristics to control for socio-economic conditions that children grow up under, since both crime and the prevalence and treatment of ADHD are highly correlated with socio-economic status (Keilow, Wu, and Obel 2020). We include age at first child (age group <25 , $25-34$, and $35+$), multipartner fertility (i.e., whether the parent has children with more than one partner), highest education (primary, secondary and tertiary education), labor force participation (employed, unemployed, outside the labor force, and in education end of year), unemployment rate (measured as % of the year unemployed), and gross income including transfers (measured in 10,000 DKK at 2015 level). We also include indicators of parental crime (convicted of violations against the Penal Code, which excludes traffic offences and narcotics misdemeanors) and measures of parental purchases of ADHD or other psychoactive medication to capture intergenerational transmission of criminality and ADHD. All parental characteristics are measured prior to the birth of the child, except from multipartner fertility (which is measured when the child turn 6)

and parental medication purchases (which is measured prior to the child turning 15 due to data availability).

RESEARCH DESIGN

To estimate potential crime-reducing effects of medical ADHD treatment, we build on the finding from the medical literature showing that treatment response is plausibly random with respect to individual, treatment, and family level characteristics. Ideally, we would want to compare crime rates between those who (for unconfounded reasons) did and did not respond well to treatment. However, without repeated observations on response to medication, we cannot directly identify medical non-response. Instead, we use the contrast between consistent and discontinued treatment as proxies for medical response to ADHD medication (similarly to Keilow, Holm, and Fallesen 2018). DTP serves as a proxy for medical non-response, and CTP serve as a proxy for successful/effective medical ADHD treatment, with ATP capturing the intermediate category. If these treatment patterns capture medical non-response and effective medical treatment without error (i.e., perfect compliance), the difference in crime between the CTP and DTP groups captures the effect of effective medical ADHD treatment. CTP could also include some “always-takers”, who consistently take the medication despite no reduction in symptoms, which would bias results downward. The presence of never-takers who would respond well to treatment but refrain from taking it irrespective of symptom reductions is unlikely because of the sample restrictions, which conditions on a) only entering medical treatment, and b) showing strong initial commitment to treatment by conditioning on minimum three medication purchases. ATP represents partial compliance. The difference between ATP and CTP thus depends on whether the average treatment pattern in this group resembles a CTP or DTP the most as well as reflecting any effects of (partial) medical treatment of ADHD.

We estimate the following model using OLS:

$$Crime_i = \alpha + \beta_1 Discontinued_i + \beta_2 Ambiguous_i + \mathbf{X}_i\boldsymbol{\gamma} + \varepsilon_i$$

where *Discontinued* and *Ambiguous* are dummy variables, $\mathbf{X}$ represents a set of covariates described above, and ε_i is the error term. Since CTP is the reference category, positive coefficients for β_1 and β_2 indicate a crime-reducing effect of effective medical ADHD treatment. However, β_1 and β_2 can only be interpreted as causal estimates under the assumption that treatment patterns conditional on $\mathbf{X}$ are as good as random and independent of the error term (ε_i). There are, however, several ways such an assumption might be violated. Below, we address each of these potential sources of bias as best as the available data allow to render plausible that observed differences can be attributed to effects of medical ADHD treatment. Further, we return to these limitations in the discussion.

First, although the analytical sample consists of exactly the population of interest, the treatment margin has changed across the period we examine with a substantial increase in the prevalence of treatment. A group of children from earlier cohorts, who would have qualified for medical treatment by current standards and would have received treatment had they been later born, will not have received treatment and thus do not appear in the sample. This would be a problem of sample selection bias, and we address this using a two-step Heckman selection model for the continuous outcomes (Bushway, Johnson, and Slocum 2007). As an exclusion restriction, we use treatment propensity in municipality of residence at age 6, which, as shown by Fallesen and Wildeman (2015) predicts the probability of being treated before age 15 (i.e., it belongs in the selection equation). It does not significantly predict crime in the sample (i.e., it does not belong in the substantive equation) after the inclusion of the covariates described above (including region of residence at age 6).

A second potential issue concerns whether treatment patterns are valid proxies for medical non-response. Families might discontinue treatment for financial reasons or because the organization of everyday life makes it difficult to follow the prescribed treatment regime. Such discontinuance may reflect a socio-economic gradient in the treatment patterns (Keilow, Holm, and Fallesen 2018). To address this concern, we conduct balancing tests to see if the DTP and CTP groups resemble each other on a series of demographic and socio-economic factors as well as factors pertaining to treatment initiation and comorbidities. If there are no systematic differences between treatment patterns on observable factors, it is plausible that there should be no differences in unobservable factors as well. We also control for these factors in the main regressions and find that they do not change the substantive results.

Third, even if treatment patterns serve as a good proxy for response to medical treatment, medical non-response might not be as good as random, and the lack of consistent predictors previously found in the literature an attribute of methodological designs, not a reflection of reality. To address this concern, we construct a placebo sample, including those that begin treatment after the outcome is measured (that is after age 20). We use this placebo sample to perform a negative control analysis (Arnold et al. 2016). If the DTP group differs systematically from the CTP group on factors such as IQ, comorbidity, symptom severity, or other factors that we expect to be correlated with crime risk, then we should also expect to find placebo effects (i.e., significantly higher crime risk prior to treatment in the DTP group compared to the CTP group), assuming that potential unobserved selection into treatment patterns is similar across the two samples.

Finally, to address the potential impact of selection on unobserved factors, we conduct a bounding exercise that calculates how large the effect of unobserved factors compared to the effect of observed factors (i.e., the selection ratio) would have to be to render the results insignificant (Oster 2019). Under the assumption that selection on observables is informative

about selection on unobservables, this bounding exercise uses both movements in coefficients and R^2 under the inclusion of covariates to estimate selection-bias adjusted coefficients. This is done under different selection ratios and an $R^{\max}$ value (which is the maximum R^2 we would expect if all explanatory factors were included) of 1.3 times the observed R^2 as suggested by Oster (2019).

RESULTS

Summary Statistics

Table 1 presents the means of covariates across (1) treatment status by age 15 in the full population of birth cohorts 1986-1996, and (2) treatment patterns in the population of children receiving medical ADHD treatment before age 15. As suspected, children who received medical ADHD treatment before age 15 differ from the non-treated children on many factors. In line with previous studies (Sayal et al. 2018; Keilow, Wu, and Obel 2020), treated children are markedly more disadvantaged than non-treated children both in terms of socio-economic status and birth outcomes. Treated children are also more likely to reside in certain regions and to be born later, which points to selection into the treatment sample that is driven by variation in treatment practices across time and regions (Fallesen and Wildeman 2015). Finally, treated children are more likely to be male and less likely to be of immigrant background, which is in line with previous findings (Sayal et al. 2018). This latter finding does not necessarily speak to difference in prevalence of ADHD between children of native Danes and children with immigrant background, but could instead be due to differences in identifying, diagnosing, and treating symptoms at the family and community level.

The three rightmost columns in Table 1 compare characteristics of the children with CTP with those with ATP and DTP, respectively (see Table A-2 for a similar comparison for the sample with conviction outcomes). Comparing CTP and ATP there are some significant

differences in characteristics – most notably that a higher share with ATP enter treatment at a relatively young age, and a lower share resides outside the capital region. Given the definition of ATP, we would expect the group to include younger children, because that simply have had more time during which their treatment pattern could become ambiguous. There is a slight indication that mothers in the ATP group have more advantaged characteristics, but differences are minor.

Comparing CTP and DTP, girls and children with immigrant background are overrepresented in the DTP group (which we address in the sensitivity analysis). Additionally, a lower share of the DTP group receives a high dosage of medication, which could indicate lower symptom severity. In terms of parental characteristics, we find very few statistically or numerically significant differences between DTP and CTP groups, and no consistent indication that one group should be more disadvantaged compared to the other, which renders the assumption that treatment discontinuation is primarily driven by medical non-response plausible.

[Table 1 about here]

Table 2 presents charge and conviction outcomes by treatment status and treatment pattern. Table 2 shows that children who ever fill a prescription for ADHD medication are roughly 2.5 times more likely to get charged or convicted. As many as 25% of children treated for ADHD have been charged with a crime prior to age 18, and 23% have been convicted of one, which underscores the overrepresentation of youths with ADHD in criminal statistics (Mohr-Jensen et al. 2019). Table 2 also preempts the main analysis, showing higher crime among the group that discontinue medical treatment of ADHD compared to those that purchase medication consistently.

[Table 2 about here.]

Main Results

We now present the main results from OLS models regressing charge and conviction measures on treatment pattern indicators and covariates. The columns in Table 3 show the main coefficients under the stepwise inclusion of covariates (from M1 to M5). Compared to consistent treatment, following a DTP by age 15 results in significantly higher risks of being convicted by age 18 and a marginally significantly higher risks of being charged ($p < 0.1$). Following a DTP also results in a significantly higher number of charges/convictions. Net of controls (as shown in column M5), the effect of DTP is a 5.3 and 8.2 percentage points increase in the risk of being charged and convicted by age 18, respectively, which represent substantial increases (21% and 36%) relative to baseline charge/conviction risks. The relative increase in the number of charges/convictions is even higher (roughly 60%). Across all measures, the inclusion of covariates moves the estimated coefficients away from zero, but the differences in coefficients across models are minor. There are only slight and inconsistent indications that following an ATP increases the charge/conviction risk, which suggests that the average treatment pattern of this group more closely resembles the CTP than the DTP.

[Table 3 about here.]

The stability of the DTP estimate highlights that our large set of control variables does not act as confounders. Instead, the findings point to a crime-reducing effects of medical treatment of ADHD for those who qualify for and respond to treatment, and thus implicates ADHD as a criminogenic factor. This is in line with previous studies of the effect of medical treatment of ADHD on crime (Dalsgaard, Nielsen, and Simonsen 2014; Lichtenstein et al. 2012; Mohr-Jensen et al. 2019), thus strengthening the conclusion that medical treatment of ADHD has significant and positive externalities through the reduction of crime.

Robustness Checks

Although our main results suggests that consistent medical treatment of ADHD lowers (detected) criminal activity, we may still have concern about confounding due to selection or endogeneity. In an extensive array of robustness checks, we examine how the main findings withstand a series of placebo, sensitivity, and bounding analyses that attempt to address to what extent the main results could be driven by unobservable differences between the groups, and to what extent they are sensitive to specifications of the analytical sample and treatment patterns.

Placebo sample. Figure 3 shows estimates for the main sample (treated by age 15) and a placebo sample (treated after age 20) after controlling for covariates. The results show no significant differences in the likelihood of being charged or convicted (nor number of charges or convictions) by age 18 across treatment patterns for those that start treatment after the outcome is measured. This finding makes it less probable that the main results are driven by unobserved differences across treatment patterns, given that any relevant unobserved factors differ across treatment patterns in a similar manner in both the main and placebo samples.³

Sensitivity to sample specification. The findings are robust to different sample specifications, in which we exclude groups that could be considered outliers because they either (1) start treatment very early (< age of 5), (2) start treatment less than one year before they turn 15 (resulting in a short window to observe their treatment pattern), (3) are among the cohorts in which very few were treated (<1990), (4) are women or (5) are immigrants or descendants of immigrants (see Figure A-1). Thus, the results are not driven by these “atypical” subgroups in the analytical sample.

Sensitivity to specification of treatment patterns. In Table A-4 we show that results are sensitive to shortening the maximum amount of time that those in the DTP group purchase medication in (from 90 to 60 days) leaving less time to dose titration. This suggests that it is

those who take a longer time for dose titration (by trying two types of drugs for example) who drive the main results. Allowing for an additional month for dose titration (from 90 to 120 days), does not change the results.

We also find that results are sensitive to narrowing the maximum gap (from 90 to 60 days) between purchases for the CTP. This adjustment cuts the share with CTP by two thirds and is indeed a very restrictive definition that few are likely to meet over a long period. Again, loosening the restriction on the maximum gap (from 90 to 120 days) do not change the results, suggesting that a considerable share originally defined as ATP behave as CTP.

Results are conditional on having purchased minimum three prescriptions (see Table A-5), and the findings thus do not replicate in samples that condition on having purchased medication only once or twice. Those who do not show a strong initial commitment to medical treatment (by purchasing only one or two prescriptions) appears to be on track to commit less crime than those who discontinue after a more serious attempt at medication (three or more prescriptions). This makes sense if those that discontinue after purchasing only one or two prescriptions do so not because of medical non-response, but because they have milder symptoms (potentially coupled with a hesitancy to commit to medical treatment). Combined with the finding that estimates are weaker if DTP is restricted to include only those who purchase medication in a shorter period (max. 60 days), results suggest the importance of conditioning on an initial commitment to treatment when defining treatment discontinuation.

Sample selection correction. To address the concern that estimates might be driven by sample selection due to compositional changes of who enters the sample of treated over time, we estimate a two-step Heckman selection model estimates (with municipality dummies as

exclusion restriction) and find that this correction changes the results minimally (as shown in Table A-6).

Potential bias from selection on unobservables. As a final attempt to address the potential bias from selection on unobservables, we present coefficients adjusted for selection bias under different assumptions about the selection ratio (δ , which refers the ratio between the effect of unobserved factors relative to the effect of observed factors) in Table A-7 and A-8. Since controlling for selection on observables move the DTP coefficients away from zero (see Table 3), this means that if selection on unobservable were to go in the same direction (column 1 and 2), correcting for unobserved selection of a similar size ($\delta = 1$) or twice the size ($\delta = 2$) would only serve to increase the estimates. The original estimates still fall within the confidence intervals of these potential bias-adjusted estimates, and the results are thus quite robust to unobserved selection that drive estimates in the same overall direction as observed selection. To render the estimated effects insignificant, selection on unobservable factors would have to go in the opposite direction of selection on observables and be of a similar magnitude ($\delta = -1$ for charges) or more than twice the size ($\delta < -2$ for convictions) as selection on all observed factors combined.

Particularly for the conviction outcomes, it seems improbable that unobserved selection of such a magnitude would exist. It is not even evident that the sign of potential unobserved selection bias would go in the opposite direction of selection on observables, which overall leads to downward bias if uncontrolled for. But we can speculate to presence of potential unobserved factors that may impact the findings. Such factors could be symptom severity and comorbidity.⁴ From Table 1, we see that those with DTP receive medication of a slightly lower strength, indicating that this group initially has lower symptoms than those treated consistently. In this case, we expect estimates to be downwardly biased, as lower symptoms would presumably be tied to lower crime risk. Difference in presence of

(unmedicated) comorbidities could also be a source of bias. Particularly, the presence of conduct disorder (CD) or oppositional defiant disorder (ODD) – which are tied to crime and other antisocial behaviors (Schilling, Walsh, and Yun 2011) – could differ across treatment patterns. Recent evidence has suggested that comorbid disorder may be slightly more prevalent among medical non-responders (Johnston et al., 2015; Kim et al., 2015; Vallejo-Valdivielso et al., 2020), which would upwardly bias the results. Following the bounding estimates, such a bias would have to be at least similar in size (if not more than twice as high) as bias arising from omitting all observable covariates (including gender, parental socioeconomic status and many more) to render the findings statistically insignificant. Additionally, that type of bias would have to be uniquely occurring in the main sample, since it does not result in placebo “effects” in the placebo sample, which buttresses our belief in our main findings.

DISCUSSION

The findings from the study suggested that effective medical treatment of ADHD reduces crime, with implications both for policy and for sociological and criminological theory. For public policy, the findings highlight a positive and highly salient externality of effective medical treatment of ADHD, as it can prevent both potential offenders from starting a criminal trajectory and spare the society and potential victims the costs of crime. Securing effective medical ADHD treatment at the right margin may thus reduce crime. Here the context for the current study (and the previous studies) is not to be disregarded. In Denmark, access to ADHD medication is rather diagnostically restrictive and is primarily used in cases with modest to severe symptoms, and this is the margin to which the findings apply. However, once prescribed ADHD medication, there are few economic barriers to obtaining medical ADHD treatment, since medication is heavily subsidized in Denmark (minimum 60% of medication costs for children are covered by the government). The absence of

economic barriers should serve to make medication more accessible to low-income and disadvantaged groups, who may thus constitute a larger share of the treated in our context than in contexts, where medication constitutes a heavier financial burden.

Second, our results have implication for theories on the medicalization of behavior viewed as unwanted or deviant. Concerns about social control exerted through medicalization are invaluable (e.g., Conrad and Bergey 2014; Conrad 1992). Yet, our study joins a small but growing literature highlighting how providing medical treatment for conditions such as ADHD may both be productive in the Foucauldian understanding of the word, and lead to a more just situation for affected individuals in the Rawlsian sense of the term. Children diagnosed with ADHD tend to face more disadvantage through their life-course (Daley et al. 2015; Keilow, Wu, and Obel 2020). Medical treatment has the potential to improve academic outcomes (Keilow, Holm, and Fallesen 2018), family stability (Fallesen and Wildeman 2015; Schermerhorn et al. 2012), and ultimately mortality (Dalsgaard et al. 2015). So, whereas the risk of overdiagnosis and -medicalization should not be ignored (Currie, Stabile, and Jones 2014; King, Jennings, and Fletcher 2014), medical treatment of ADHD still has the potential of being a net-gain for both society and affected individuals.

Last, the findings has implication for criminological theory by highlighting ADHD as a relevant criminogenic factor, underscoring the salience of biological/genetic explanations for criminal behavior (Wright and Beaver 2005). ADHD is largely genetically informed (Faraone and Larsson 2019), and the medical treatment works by addressing neurological imbalances (Schilling, Walsh, and Yun 2011). That treatment appears to reduce crime speaks to a biological component of the propensity to engage in criminal behavior (unless the entire effect is driven by a reduction in adverse environmental responses to the child). Prior studies have suggested that the criminogenic influence of ADHD goes through low self-control (Unnever, Cullen, and Pratt 2003). If this is the case, then this casts doubt on the proposition

made by Hirschi and Gottfredson that ineffective parental management serves as *the* cause of low self-control. Instead, a criminogenic influence of ADHD that goes through low self-control would suggest that this commonly acknowledged driver of criminal propensity has its roots in genetic factors and biology in addition to ineffective parental management (Unnever, Cullen, and Pratt 2003; Wright and Beaver 2005).

However, even though effective medical ADHD treatment appears to have substantial crime reducing effects, this does not bring down the charge and conviction level to resemble children who presumably do not have ADHD. Even consistently and effectively treated children with ADHD still have much higher rates of charges and convictions compared to children without ADHD. This could be because medication does not remove the symptoms completely, because criminogenic set-backs have already occurred, but most likely also because children with ADHD are disadvantaged in several other areas that increase their risk of engaging in crime. Medicated or not, children with ADHD might still grow up in more disadvantaged communities and families. Thus, on average, these children may have few prospects of succeeding in conventional ways, or grow up with increased risk parental substance abuse, mental illness, and crime – all of which matter for future crime risk, and for which medication does not provide a quick fix.

Limitations

Despite the study employing extraordinarily rich data on medication patterns, crime, and family characteristics, three data limitations are noteworthy and may have implications for the interpretation of the results. First, the only available measures of crime in the administrative data are recorded encounters with the police and criminal justice system. Thus, we have no indication of any criminal behavior that goes undetected, and the current findings come with the caveat, that they might reflect differences in detection in addition to

differences in behavior. In other words, it could be that effective ADHD medication made for better criminals by making them better at covering their tracks and avoiding detection (and charges and convictions). Given the size of the estimates it is however unlikely that this can explain all the differences in observed crime risk.

Second, since medical non-response is unobserved in the data, we let those following a discontinued treatment pattern serve as proxies for medical non-responders. Being imperfect proxies, they could also include individuals who discontinue for other reasons than medical non-response. For example, we might worry that discontinuing treatment might be more prevalent among more disadvantaged families (Keilow, Holm, and Fallesen 2018). However, given the observed similarity with those following a consistent treatment pattern across a wide range of characteristics, there is little to suggest that group with a discontinued treatment pattern are different from those that continue treatment – other than in their response to medication.

Third, the available data does not allow for a good grasp of the presence of co-morbidity and symptom severity, which could be cofounders. Particularly, the presence of conduct disorder (CD) or oppositional defiant disorder (ODD) could differ across treatment patterns, and (upwardly) bias our results. However, if this were the case and if this type of selection persisted across timing of treatment, then we would also expect to see significant differences across treatment patterns in the placebo sample, which is not the case.

Future Studies

With the results from this study, we can now be fairly confident that effective ADHD medication reduces crime (measured as charges and convictions) among the average child prescribed medication. Future studies should examine which pathways effective ADHD medication reduces crime through. ADHD medication has consistently been shown to

alleviate ADHD symptoms (Sayal et al. 2018), but is the reduction in crime a direct effect of the symptom alleviation or is it mediated by an improvement in how the child functions and is treated in the home or school environment? The answer to this question has implications for how closely tied ADHD behavior and crime is, and it is a question ripe for future exploration.

CONCLUSION

We found that compared to children with a consistent treatment pattern, a pattern of discontinued ADHD treatment resulted in substantial increase in both the risk of being charged or convicted and the number of charges and convictions by age 18, net of a wide range of demographic, treatment specific, and socio-economic characteristics. The finding applied to children, who started medical ADHD treatment prior to the age of 15. To reach this conclusion, we utilized the fact that medical non-response among those treated for ADHD occurred in a plausibly random manner. Using individual-level longitudinal data on medication purchases, we identified a group of children who most likely discontinued treatment due to medical non-response and compared this group to children who followed a consistent treatment pattern, while controlling for a wide range of covariates. The results were robust to different sample specifications, various adjustments for potential selection bias, and placebo tests suggesting that effects are not driven by inherent differences between treatment patterns unaccounted for in this study. Combined with the findings from prior studies (Lichtenstein et al. 2012; Mohr-Jensen et al. 2019; Dalsgaard, Nielsen, and Simonsen 2014) the robustness and size of the estimated effects suggested that ADHD medication reduced crime among children prescribed medication.

REFERENCES:

- Andersen, Signe Hald, Peter Fallesen, Jeppe Elholm Madsen, and Christopher Wildeman. 2021. "Paternal Incarceration and Children's Use of Psychotropic Medication."
- Arnold, Benjamin F., Ayse Ercumen, Jade Benjamin-Chung, and John M. Colford. 2016. "Brief Report: Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies." *Epidemiology* 27 (5): 637–41.
- Barkely, Russel A. 2015. "Etiologies of ADHD." In *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment.*, edited by Russell A. Barkley, 4th ed., 356–90. Guilford Press.
- Bushway, Shawn D., Brian D Johnson, and Lee Ann Slocum. 2007. "Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology." *Journal of Quantitative Criminology* 23 (2): 151–78.
- Chazan, Rodrigo, Carolina Borowski, Thiago Pianca, Henrique Ludwig, Luis Augusto Rohde, and Guilherme Polanczyk. 2011. "Do Phenotypic Characteristics, Parental Psychopathology, Family Functioning, and Environmental Stressors Have a Role in the Response to Methylphenidate in Children with Attention-Deficit/Hyperactivity Disorder?: A Naturalistic Study from a Developing Coun." *Journal of Clinical Psychopharmacology* 31 (3): 309–17.
- Conrad, Peter. 1992. "Medicalization and Social Control." *Annual Review of Sociology* 18 (1): 209–32.
- . 2005. "The Shifting Engines of Medicalization." *Journal of Health and Social Behavior* 46 (1): 3–14.
- Conrad, Peter, and Meredith R Bergey. 2014. "The Impending Globalization of ADHD: Notes on the Expansion and Growth of a Medicalized Disorder." *Social Science and Medicine* 122: 31–43.
- Currie, Janet, Mark Stabile, and Lauren Jones. 2014. "Do Stimulant Medications Improve Educational and Behavioral Outcomes for Children with ADHD?" *Journal of Health Economics* 37 (1): 58–69.
- Daley, David, Rasmus Højbjerg Jacobsen, Anne-Mette Lange, Anders Sørensen, and Jeanette Walldorf. 2015. *Costing Adult Attention Deficit Hyperactivity Disorder*. Oxford

University Press. Oxford University Press.

- Dalsgaard, Søren, Maria Knøth Humlum, Helena Skyt Nielsen, and Marianne Simonsen. 2014. "Common Danish Standards in Prescribing Medication for Children and Adolescents with ADHD." *European Child and Adolescent Psychiatry* 23 (9): 841–44.
- Dalsgaard, Søren, Preben Bo Mortensen, Morten Frydenberg, and Per Hove Thomsen. 2013. "Long-Term Criminal Outcome of Children with Attention Deficit Hyperactivity Disorder." *Criminal Behaviour and Mental Health* 23 (2): 86–98.
- Dalsgaard, Søren, Helena Skyt Nielsen, and Marianne Simonsen. 2013. "Five-Fold Increase in National Prevalence Rates of Attention-Deficit/ Hyperactivity Disorder Medications for Children and Adolescents with Autism Spectrum Disorder, Attention-Deficit/Hyperactivity Disorder, and Other Psychiatric Disorders: A Danish Register-Based Study." *Journal of Child and Adolescent Psychopharmacology* 23 (7): 432–39.
- . 2014. "Consequences of ADHD Medication Use for Children's Outcomes." *Journal of Health Economics* 37 (1): 137–51.
- Dalsgaard, Søren, Søren Dinesen Østergaard, James F. Leckman, Preben Bo Mortensen, and Marianne Giørtz Pedersen. 2015. "Mortality in Children, Adolescents, and Adults with Attention Deficit Hyperactivity Disorder: A Nationwide Cohort Study." *The Lancet* 385 (9983): 2190–96.
- Danish Health Authority. 2014. "National Klinisk Retningslinje for Udredning Og Behandling Af ADHD Hos Børn Og Unge."
- DuPaul, George J, and Joshua M Langberg. 2015. "Educational Impairments in Children with ADHD." In *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*, edited by Russel Barkely, 4th ed., 169–90. Guilford Publications.
- Elder, Todd E. 2010. "The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates." *Journal of Health Economics* 29: 641–56.
- Fallesen, Peter, and Christopher Wildeman. 2015. "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* 56 (3): 398–414.
- Faraone, Stephen V., Tobias Banaschewski, David Coghill, Yi Zheng, Joseph Biederman, Mark A. Bellgrove, Jeffrey H. Newcorn, et al. 2021. "The World Federation of ADHD

- International Consensus Statement: 208 Evidence-Based Conclusions about the Disorder.” *Neuroscience & Biobehavioral Reviews*, February.
- Faraone, Stephen V, and Henrik Larsson. 2019. “Genetics of Attention Deficit Hyperactivity Disorder.” *Molecular Psychiatry* 24: 562–75.
- Gottfredson, Michael R., and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford, California: Stanford University Press.
- Gray, Jeremy R, and Jerome Kagan. 2000. “The Challenge of Predicting Which Children with Attention Deficit-Hyperactivity Disorder Will Respond Positively to Methylphenidate.” *Journal of Applied Developmental Psychology* 21 (5): 471–89.
- Johnston, Blair A., David Coghill, Keith Matthews, and J. Douglas Steele. 2015. “Predicting Methylphenidate Response in Attention Deficit Hyperactivity Disorder: A Preliminary Study.” *Journal of Psychopharmacology* 29 (1): 24–30.
- Johnston, Charlotte, and Andrea Chronis-Rusciano. 2015. “Families and ADHD.” In *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*, edited by Russel Barkely, 4th ed., 191–209. Guilford Publications.
- Keilow, Maria, Anders Holm, and Peter Fallesen. 2018. “Medical Treatment of Attention Deficit/ Hyperactivity Disorder (ADHD) and Children’s Academic Performance.” Edited by Nouchine Hadjikhani. *PLoS ONE* 13 (11): e0207905.
- Keilow, Maria, Chunsen Wu, and Carsten Obel. 2020. “Cumulative Social Disadvantage and Risk of Attention Deficit Hyperactivity Disorder: Results from a Nationwide Cohort Study.” *SSM - Population Health* 10 (April): 100548.
- Kim, Jae-Won, Vinod Sharma, and Neal D Ryan. 2015. “Predicting Methylphenidate Response in ADHD Using Machine Learning Approaches.” *International Journal of Neuropsychopharmacology*, 1–7.
- King, Marissa D, Jennifer Jennings, and Jason M Fletcher. 2014. “Medical Adaptation to Academic Pressure: Schooling, Stimulant Use, and Socioeconomic Status.” *American Sociological Review* 79 (6): 1039–66.
- Kvist, Anette Primdal, Helena Skyt Nielsen, and Marianne Simonsen. 2013. “The Importance of Children’s ADHD for Parents’ Relationship Stability and Labor Supply.” *Social Science and Medicine* 88: 30–38.

- Lichtenstein, Paul, Linda Halldner, Johan Zetterqvist, Arvid Sjölander, Eva Serlachius, Seena Fazel, Niklas Långström, and Henrik Larsson. 2012. "Medication for Attention Deficit–Hyperactivity Disorder and Criminality." *New England Journal of Medicine* 367 (21): 2006–2014.
- McQuade, Julia D., and Betsy Hoza. 2015. "Peer Relationships of Children with ADHD." In *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment.*, edited by Russel Barkely, 4th ed., 210–22. Guilford Publications.
- Mohr-Jensen, Christina, Charlotte Müller Bisgaard, Søren Kjærsgaard Boldsen, and Hans-Christoph Steinhausen. 2019. "Attention-Deficit/Hyperactivity Disorder in Childhood and Adolescence and the Risk of Crime in Young Adulthood in a Danish Nationwide Study." *Journal of the American Academy of Child & Adolescent Psychiatry*, February.
- Mohr-Jensen, Christina, and Hans Christoph Steinhausen. 2016. "A Meta-Analysis and Systematic Review of the Risks Associated with Childhood Attention-Deficit Hyperactivity Disorder on Long-Term Outcome of Arrests, Convictions, and Incarcerations." *Clinical Psychology Review*.
- NICE. 2018. "Attention Deficit Hyperactivity Disorder: Diagnosis and Management (NICE Guideline)." *NICE - National Institute for Health and Care Excellence*.
- Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence" 37 (2): 187–204.
- Pelham, William E., and Bradley Smith. 2000. "Prediction and Measurement of Individual Responses to Ritalin by Children and Adolescents with Attention Deficit Hyperactivity Disorder." In *Ritalin, Theory and Practice*, edited by Laurence L Greenhill and Betty B Osman, Second, 193–218. Larchmont, NY: Mary Ann Liebert Publishers.
- Perrild, Bjørn. 2020. "ADHD, Børn Og Unge." Sundhed.Dk. 2020.
- Polanczyk, Guilherme, Maurício Silva De Lima, Bernardo Lessa Horta, Joseph Biederman, and Luis Augusto Rohde. 2007. "The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis." *American Journal of Psychiatry* 164 (6): 942–48.
- Pratt, Travis C, Francis T Cullen, Kristie R Blevins, Leah Daigle, and James D Unnever. 2002. "The Relationship of Attention Deficit Hyperactivity Disorder to Crime and Delinquency: A Meta-Analysis." *International Journal of Police Science &*

Management 4 (4): 344–60.

Region Midtjylland. 2016. “Information Om Medicin Mod ADHD Til Børn Og Unge.”

Roberts, Walter, Richard Milich, and Russell A. Barkley. 2015. “Primary Symptoms, Diagnostic Criteria, Subtyping, and Prevalence of ADHD.” In *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment.*, edited by Russel A. Barkely, 4th ed., 51–80. Guilford Publications.

Sayal, Kapil, Vibhore Prasad, David Daley, Tamsin Ford, and David Coghill. 2018. “ADHD in Children and Young People: Prevalence, Care Pathways, and Service Provision.” *The Lancet Psychiatry*.

Schermerhorn, Alice C., Brian M. DOnofrio, Wendy S. Slutske, Robert E. Emery, Eric Turkheimer, K. Paige Harden, Andrew C. Heath, and Nicholas 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.

Schilling, Catrina M, Anthony Walsh, and Ilhong Yun. 2011. “ADHD and Criminality: A Primer on the Genetic, Neurobiological, Evolutionary, and Treatment Literature for Criminologists.” *Journal of Criminal Justice*.

Sonneville, Leo M.J. de, Charles Njioiktjien, and Hannie Bos. 1994. “Methylphenidate and Information Processing. Part 1: Differentiation between Responders and Nonresponders; Part 2: Efficacy in Responders.” *Journal of Clinical and Experimental Neuropsychology* 16 (6): 877–97.

Unnever, James D, Francis T Cullen, and Travis C Pratt. 2003. “Parental Management, ADHD, and Delinquent Involvement: Reassessing Gottfredson and Hirschi’s General Theory.” *Justice Quarterly*.

Vallejo-Valdivielso, María, Pilar de Castro-Manglano, Azucena Díez-Suárez, Juan J Marín-Méndez, and Cesar A Soutullo. 2020. “Clinical and Neuropsychological Predictors of Methylphenidate Response in Children and Adolescents with ADHD: A Naturalistic Follow-up Study in a Spanish Sample.” *Clinical Practice & Epidemiology in Mental Health* 15 (1): 160–71.

Vaughan, Brigitte S, John S March, and Christopher J Kratochvil. 2012. “The Evidence-Based Pharmacological Treatment of Paediatric ADHD.” *International Journal of*

Neuropsychopharmacology 15 (1): 27–39.

- Wallach-Kildemoes, Helle, Anne M. Skovgaard, Karsten Thielen, Anton Pottegård, and Laust H. Mortensen. 2015. “Social Adversity and Regional Differences in Prescribing of ADHD Medication for School-Age Children.” *Journal of Developmental and Behavioral Pediatrics* 36 (5): 330–41.
- Wright, John Paul, and Kevin M. Beaver. 2005. “Do Parents Matter in Creating Self-Control in Their Children? A Genetically Informed Test of Gottfredson and Hirschi’s Theory of Low Self-Control.” *Criminology* 43 (4): 1169–1202.
- Young, S, D Moss, O Sedgwick, M Fridman, and P Hodgkins. 2015. “A Meta-Analysis of the Prevalence of Attention Deficit Hyperactivity Disorder in Incarcerated Populations.” *Psychological Medicine* 45: 247–58.
- Zoëga, Helga, Unnur A. Valdimarsdóttir, and Sonia Hernández-Díaz. 2012. “Age, Academic Performance, and Stimulant Prescribing for ADHD: A Nationwide Cohort Study.” *Pediatrics* 130 (6): 1012–18.

ENDNOTE

¹ There is high variability in the share of children diagnosed or treated for ADHD as diagnostic and treatment procedures vary across countries and over time (Polanczyk et al. 2007; Sayal et al. 2018). However, the differences in diagnose prevalence are largely explained by differences in diagnostic criteria and treatment regimens (Polanczyk et al., 2007) and increasing convergence could be attributable to spread of the DSM diagnostic criteria.

² The few supported environmental explanations are biological rather than social in nature (i.e., early exposure to lead, maternal smoking and alcohol consumption during pregnancy, birth complications and traumatic brain injury, Barkely 2015). Recent work has however provided indications that social stressor, such a paternal incarceration, may increase likelihood of starting medical treatment (Andersen et al. 2021).

³ Naturally, the two samples differ in that the placebo sample draws more heavily on early cohorts and on youth from regions that were late to diagnosing and treating children for ADHD. The placebo sample also has a higher crime rate, which is to be expected since they have been unmedicated during the observation period. Additionally, the placebo sample appears to be of lower socio-economic status, which suggests that being late to treatment is more likely among disadvantaged youths. When comparing characteristics across treatment patterns in the placebo sample (see Table A-3), there is some evidence pointing in the direction that those who discontinue are more disadvantaged than those who continue treatment (which was not evident in the main sample), and women are underrepresented in the group with DTP (which contrasts with the main sample). These two features suggest that selection on observables into treatment patterns differ across main and placebo samples, and treatment pattern in adulthood may thus be more determined by social factors than by

medical response compared to treatment patterns in childhood. However, controlling for these covariates there are no significant differences in outcomes between treatment patterns in the placebo sample.

⁴ We attempt to account for both symptom severity and comorbidity (by strength of medication and treatment for other psychiatric disorders), but the measures are far from perfect.

Figure 1. Treatment patterns by age 15

Source: Own calculations based on data from Statistics Denmark.

Note: Plots are for all children treated before age 15 from birth cohorts 1986-1996 and who fill minimum three prescriptions. Values in plot a) and b) are conditional on being within the observation window in a given month (that is N for month 1 and month 24 are not the same). Similarly values in plot c) and d) are conditional on having a next prescription within the observation window.

Figure 2. Coefficients from main and placebo samples

Source: Own calculations based on data from Statistics Denmark.

Note: Placebo sample is defined as those who commence treatment at age 20 or later (i.e., after the outcome is measured). Lines indicate 95% confidence intervals.

Table 1. Summary statistics by treatment status and pattern

	Full population		Treatment pattern by age 15		
	Not treated	Treated by age 15	Consistent	Ambiguous	Discontinued
Child characteristics					
Cohort 1986	0.079	0.018***	0.019	0.016	0.019
Cohort 1987	0.080	0.024***	0.027	0.021	0.031
Cohort 1988	0.084	0.028***	0.035	0.026+	0.028
Cohort 1989	0.088	0.043***	0.043	0.036	0.040
Cohort 1990	0.091	0.049***	0.045	0.048	0.025
Cohort 1991	0.092	0.069***	0.059	0.069	0.093*
Cohort 1992	0.097	0.094	0.079	0.094+	0.102
Cohort 1993	0.096	0.118***	0.119	0.115	0.152
Cohort 1994	0.099	0.157***	0.159	0.157	0.152
Cohort 1995	0.098	0.192***	0.198	0.198	0.207
Cohort 1996	0.095	0.208***	0.216	0.219	0.152*
Female	0.490	0.191***	0.186	0.172	0.272***
Immigrant background	0.077	0.032***	0.024	0.027	0.053**
Premature	0.050	0.084***	0.077	0.085	0.093
Missing Apgar score	0.049	0.052	0.049	0.051	0.050
Low Apgar score	0.010	0.016***	0.014	0.016	-
Missing gestational age	0.051	0.057*	0.056	0.056	0.053
Low birth weight	0.046	0.076***	0.070	0.078	0.059
Missing birth weight	0.047	0.051+	0.048	0.05	0.046
Region Sjælland	0.148	0.138**	0.152	0.126*	0.183
Region Syddanmark	0.253	0.216***	0.264	0.208***	0.180**
Region Midtjylland	0.178	0.192***	0.220	0.195*	0.170*
Region Nordjylland	0.142	0.133*	0.145	0.131	0.139
Lives w. both	0.788	0.605***	0.604	0.605	0.613
Lives w. single parent	0.153	0.260***	0.254	0.258	0.266
Lives w. parent and step	0.052	0.103***	0.116	0.102	0.093
Lives w. other	0.006	0.033***	0.027	0.034	0.028
Missing lives. w.	0.001	-	-	-	-
Other psychoactive med.	0.028	0.247***	0.252	0.245	0.251
Treatment characteristics					
High dosage(first)			0.167	0.188+	0.108**
High dosage(third)			0.121	0.216***	0.127
Young at treatment			0.466	0.703***	0.486
Old at treatment			0.08	0.012***	0.053+
Multiple prescriptions (first)			0.064	0.035***	0.046
Multiple prescriptions (third)			0.103	0.095	0.087
Father characteristics					
Aged <25 at first child	0.264	0.317***	0.330	0.320	0.300
Aged 35+ at first child	0.106	0.116**	0.104	0.115	0.121
Multi-partner fertility	0.123	0.182***	0.179	0.179	0.176
Not in FAIN	0.012	0.018***	0.014	0.019	-
Missing father	0.017	0.024***	0.021	0.023	0.025
Upper secondary educ.	0.184	0.202***	0.202	0.203	0.180
Tertiary educ.	0.645	0.536***	0.514	0.539	0.520
Missing educ.	0.063	0.057*	0.054	0.053	0.059

Unemployed	0.071	0.107***	0.094	0.108	0.133*
Outside labor force	0.050	0.071***	0.072	0.070	0.062
In education	0.007	0.004***	-	0.004	-
Missing employment	0.032	0.021***	0.019	0.021	0.022
% unemployed	8.286	12.074***	12.105	11.982	13.618
Gross income 10.000 DKK	32.679	29.860***	29.300	29.986	29.078
Penal conviction 5 years PB	0.073	0.142***	0.148	0.140	0.158
ADHD medication	0.000	0.001***	-	0.001	-
Other psychoactive med.	0.069	0.149***	0.165	0.148	0.118*
Mother characteristics					
Aged <25 at first child	0.467	0.555***	0.609	0.543***	0.557+
Aged 35+ at first child	0.029	0.033*	0.029	0.032	0.050+
Multi-partner fertility	0.126	0.214***	0.208	0.213	0.214
Not in FAIN	0.002	0.004**	-	0.004	-
Missing mother	0.001	-	-	-	-
Upper secondary educ.	0.312	0.373***	0.378	0.370	0.362
Tertiary educ.	0.556	0.436***	0.405	0.440*	0.443
Missing educ.	0.061	0.039***	0.037	0.035	0.050
Unemployed	0.135	0.170***	0.171	0.166	0.198
Outside labor force	0.105	0.158***	0.174	0.156	0.167
In education	0.011	0.011	0.007	0.011	-
Missing employment	0.039	0.024***	0.026	0.022	0.025
% unemployed	13.855	17.485***	18.094	17.222	17.734
Gross income 10.000 DKK	21.696	20.437***	19.500	20.664***	19.991
Penal conviction 5 years PB	0.024	0.051***	0.058	0.048	0.046
ADHD medication	0.000	0.004***	-	0.005	-
Other psychoactive med.	0.097	0.219***	0.229	0.216	0.192
Observations	707138	8604	1244	5823	323

Source: Own calculations based on data from Statistics Denmark.

Note: Stars indicates significant differences in means between a) those treated with ADHD medication and not in the full population, and b) those with consistent treatment patterns and ambiguous/discontinued treatment patterns. * p<0.05, ** p<0.01, *** p<0.001. PB = pre birth of the child. "-" indicates clouded cell because cell size is below 5 observations.

Table 2. Mean outcomes by treatment status and pattern

	Full population		Treatment pattern by age 15		
	Not treated	Treated by age 15	Consistent	Ambiguous	Discontinued
Charges					
Charged by age 18	0.104	0.255***	0.258	0.253	0.294
# charges by age 18	0.221	0.801***	0.830	0.769	1.238*
Observations	707138	8604	1244	5823	323
Convictions					
Convicted by age 18	0.093	0.230***	0.222	0.232	0.295*
# convictions by age 18	0.136	0.403***	0.390	0.400	0.604**
Observations	570492	5167	729	3396	207

Source: Own calculations based on data from Statistics Denmark.

Note: Stars indicates significant differences in means between a) those treated with ADHD medication and not in the full population, and b) those with consistent treatment patterns and ambiguous/discontinued treatment patterns. + p<0.01 * p<0.05, ** p<0.01, *** p<0.001

Table 3. Estimation results from OLS models

	M1	M2	M3	M4	M5
Charged by age 18					
Ambiguous	-0.005 (0.014)	0.002 (0.014)	0.02 (0.014)	0.025+ (0.014)	0.024+ (0.014)
Discontinued	0.036 (0.028)	0.038 (0.028)	0.051+ (0.028)	0.054+ (0.028)	0.053+ (0.028)
Observations	7390	7390	7390	7390	7390
Mean	0.255	0.255	0.255	0.255	0.255
# of charges by age 18					
Ambiguous	-0.062 (0.089)	-0.045 (0.092)	-0.008 (0.089)	0.014 (0.089)	0.009 (0.089)
Discontinued	0.408* (0.205)	0.408* (0.205)	0.455* (0.201)	0.465* (0.201)	0.457* (0.201)
Observations	7390	7390	7390	7390	7390
Mean	0.800	0.800	0.800	0.800	0.800
Convicted by age 18					
Ambiguous	0.01 (0.017)	0.019 (0.017)	0.031+ (0.017)	0.030+ (0.017)	0.029+ (0.017)
Discontinued	0.072* (0.035)	0.072* (0.035)	0.089* (0.035)	0.083* (0.035)	0.082* (0.035)
Observations	4332	4332	4332	4332	4332
Mean	0.233	0.233	0.233	0.233	0.233
# convictions by age 18					
Ambiguous	0.01 (0.038)	0.03 (0.038)	0.051 (0.037)	0.049 (0.037)	0.046 (0.037)
Discontinued	0.214* (0.091)	0.213* (0.090)	0.251** (0.090)	0.236** (0.090)	0.233** (0.090)
Observations	4332	4332	4332	4332	4332
Mean	0.408	0.408	0.408	0.408	0.408
Treatment					
covariates		X	X	X	X
Child demographics			X	X	X
Other psychoactive med.			X	X	X
Parental demographics				X	X
Parental SES				X	X
Parental psychoactive med.					X

Source: Own calculations based on data from Statistics Denmark.

Note: The table reports the coefficients for ambiguous and discontinued treatment patterns (relative to consistent) under different set of controls. Standard errors are robust. + p<0.1,

* p<0.05, ** p<0.01, *** p<0.001

SUPPLEMENTARY TABLES AND FIGURES

Table A-1. Stepwise sample selection

	Step 1	Step 2	Step 3	Step 4
Count	715742	23494	8218	7390
Select variables, mean				
Cohort	1991.24	1991.537	1993.202	1993.291
Male	0.513	0.675	0.817	0.822
Consistent treatment		0.180	0.066	0.073
Ambiguous treatment		0.647	0.843	0.906
Discontinued treatment		0.173	0.092	0.021
Consistent treatment by age 15			0.151	0.168
Ambiguous treatment by age 15			0.716	0.788
Discontinued treatment by age 15			0.132	0.044
Added sample criteria:				
Treated		X	X	X
Treated by age 15 (min. 90 days prior)			X	X
Min. 3 prescriptions				X

Source: Own calculations based on data from Statistics Denmark.

Note: Table shows distribution of treatment patterns and select covariates under stepwise sample selection.

Table A-2. Summary statistics by treatment status and pattern. Conviction sample.

	Full population		Treatment pattern by age 15		
	Not treated	Treated by age 15	Consistent	Ambiguous	Discontinued
Child characteristics					
Cohort 1986	0.098	0.030***	0.033	0.028	0.029
Cohort 1987	0.100	0.040***	0.047	0.036	0.048
Cohort 1988	0.104	0.047***	0.060	0.044+	0.043
Cohort 1989	0.109	0.071***	0.073	0.062	0.063
Cohort 1990	0.113	0.081***	0.077	0.083	0.039+
Cohort 1991	0.115	0.115	0.102	0.117	0.145+
Cohort 1992	0.120	0.157***	0.134	0.162+	0.159
Cohort 1993	0.119	0.197***	0.203	0.198	0.237
Cohort 1994	0.123	0.261***	0.272	0.270	0.237
Cohort 1995	-	-	-	-	-
Cohort 1996	-	-	-	-	-
Female	0.490	0.172***	0.158	0.151	0.256**
Immigrant background	0.073	0.028***	0.030	0.023	0.029
Premature	0.049	0.082***	0.069	0.084	0.087
Missing Apgar score	0.050	0.054	0.053	0.051	0.063
Low Apgar score	0.009	0.019***	0.016	0.020	-
Missing gestational age	0.051	0.058*	0.056	0.055	0.068
Low birth weight	0.046	0.080***	0.073	0.082	0.058
Missing birth weight	0.047	0.052+	0.052	0.048	0.053
Region Hovedstaden	0.278	0.322***	0.226	0.344***	0.333**
Region Sjælland	0.147	0.124***	0.132	0.110+	0.203*
Region Syddanmark	0.254	0.225***	0.288	0.210***	0.188**
Region Midtjylland	0.178	0.207***	0.230	0.214	0.145**
Region Nordjylland	0.143	0.122***	0.123	0.122	0.130
Lives w. both (age 6)	0.791	0.609***	0.628	0.605	0.628
Lives w. single parent	0.153	0.264***	0.239	0.268	0.227
Lives w. parent and step	0.048	0.094***	0.107	0.094	0.121
Lives w. other	0.006	0.032***	0.026	0.033	0.024
Missing lives. w.	0.001	-	-	-	-
Other psychoactive med.	0.023	0.224***	0.214	0.220	0.232
Treatment characteristics					
High dosage(first)			0.163	0.189	0.106*
High dosage(third)			0.093	0.213***	0.135+
Young at treatment			0.645	0.855***	0.652
Old at treatment			-	-	-
Multiple prescriptions (first)			0.045	0.029*	0.043
Multiple prescr. (third)			0.077	0.076	0.068
Father characteristics					
Aged <25 at first child	0.276	0.333***	0.355	0.336	0.324
Aged 35+ at first child	0.102	0.114**	0.106	0.107	0.121
Multi-partner fertility	0.125	0.187***	0.173	0.186	0.213
Not in FAIN	0.012	0.017***	0.012	0.017	-
Missing father	0.017	0.026***	0.023	0.024	0.029
Upper secondary educ.	0.182	0.194*	0.200	0.193	0.159
Tertiary educ.	0.641	0.535***	0.510	0.540	0.527
Missing educ.	0.064	0.061	0.048	0.059	0.058

Unemployed	0.073	0.121***	0.110	0.121	0.135
Outside labor force	0.048	0.067***	0.073	0.067	0.063
In education	0.006	0.002***	-	0.002***	-
Missing employment	0.032	0.023***	0.022	0.021	0.024
% unemployed	8.229	12.561***	12.783	12.365	14.041
Gross income 10.000 DKK	32.544	29.247***	28.126	29.522+	28.973
Penal conviction 5 y. PB	0.074	0.150***	0.150	0.150	0.193
ADHD medication	0.000	0.002***	-	0.002***	-
Other psychoactive med.	0.053	0.121***	0.123	0.122	0.116
Mother characteristics					
Aged <25 at first child	0.484	0.573***	0.635	0.562***	0.589
Aged 35+ at first child	0.027	0.033**	0.021	0.031	0.048*
Multi-partner fertility	0.127	0.216***	0.189	0.218+	0.251+
Not in FAIN	0.002	0.004**	-	0.005***	-
Missing mother	0.001	-	-	-	-
Upper secondary educ.	0.316	0.376***	0.398	0.371	0.357
Tertiary educ.	0.546	0.423***	0.377	0.427*	0.425
Missing educ.	0.062	0.044***	0.044	0.037	0.058
Unemployed	0.139	0.185***	0.187	0.181	0.242+
Outside labor force	0.098	0.143***	0.173	0.136**	0.169
In education	0.009	0.005**	-	0.005***	-
Missing employment	0.038	0.026***	0.033	0.021+	0.034
% unemployed	14.171	18.473***	18.816	18.097	20.614
Gross income 10.000 DKK	21.430	19.709***	18.313	20.038***	18.092
Penal conviction 5 y. PB	0.024	0.056***	0.062	0.052	0.058
ADHD medication	0.000	0.005***	-	0.006***	-
Other psychoactive med.	0.078	0.186***	0.180	0.183	0.159
Observations	570492	5167	729	3396	207

Source: Own calculations based on data from Statistics Denmark.

Note: Stars indicates significant differences in means between a) those treated with ADHD medication and not in the full population, and b) those with consistent treatment patterns and ambiguous/discontinued treatment patterns. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. PB = pre birth of the child. "-" indicates clouded cell because cell size is below 5 observations.

Table A-3. Summary statistics by treatment pattern after age 20 (placebo)

	Treatment pattern after age 20		
	Consistent	Ambiguous	Discontinued
Child characteristics			
Cohort 1986	0.168	0.205**	0.161
Cohort 1987	0.168	0.212***	0.126*
Cohort 1988	0.174	0.209**	0.164
Cohort 1989	0.152	0.172+	0.152
Cohort 1990	0.158	0.111***	0.137
Cohort 1991	0.103	0.060***	0.127+
Cohort 1992	0.062	0.029***	0.090*
Cohort 1993	0.014	-	0.042***
Cohort 1994	-	-	-
Cohort 1995	-	-	-
Cohort 1996	-	-	-
Female	0.442	0.378***	0.402+
Immigrant background	0.020	0.025	0.020
Premature	0.068	0.056	0.062
Missing Apgar score	0.037	0.045	0.037
Low Apgar score	0.007	0.009	0.012
Missing gestational age	0.032	0.044*	0.045
Low birth weight	0.069	0.066	0.065
Missing birth weight	0.028	0.039+	0.034
Region Sjælland	0.141	0.164*	0.183*
Region Syddanmark	0.247	0.205**	0.235
Region Midtjylland	0.238	0.236	0.199*
Region Nordjylland	0.168	0.168	0.151
Lives w. both (age 6)	0.659	0.621*	0.630
Lives w. single parent (age 6)	0.247	0.280*	0.271
Lives w. parent and step (age 6)	0.082	0.081	0.074
Lives w. other (age 6)	0.012	0.019	0.025*
Missing lives. w. (age 6)	-	-	-
Treatment characteristics			
High dosage(first)	0.147	0.176*	0.112*
High dosage(third)	0.198	0.231*	0.166+
Young at treatment for cohort	0.011	0.036***	0.017
Old at treatment for cohort	0.690	0.378***	0.754**
Multiple prescriptions (first)	0.057	0.047	0.040
Multiple prescriptions (third)	0.106	0.118	0.126
Father characteristics			
Aged <25 at first child	0.379	0.416*	0.394
Aged 35+ at first child	0.102	0.097	0.099
Multi-partner fertility	0.197	0.193	0.204
Not in FAIN	0.023	0.022	0.017
Missing father	0.021	0.021	0.018
Upper secondary educ.	0.188	0.172	0.149*
Tertiary educ.	0.534	0.509+	0.501
Missing educ.	0.044	0.065**	0.070*

Unemployed	0.089	0.102	0.126*
Outside labor force	0.049	0.067*	0.064
In education	0.006	0.009	-
Missing employment	0.013	0.021*	0.030**
% unemployed	10.999	11.989	13.352*
Gross income 10.000 DKK	30.03	29.054+	28.665+
Penal conviction 5 years PB	0.135	0.174***	0.174*
ADHD medication	-	-	-
Other psychoactive medication	0.030	0.024	0.054**
Mother characteristics			
Aged <25 at first child	0.632	0.663*	0.620
Aged 35+ at first child	0.021	0.015	0.023
Multi-partner fertility	0.220	0.221	0.213
Not in FAIN	0.004	-	-
Missing mother	-	-	-
Upper secondary educ.	0.363	0.370	0.370
Tertiary educ.	0.365	0.358	0.360
Missing educ.	0.032	0.043+	0.049+
Unemployed	0.198	0.191	0.228
Outside labor force	0.128	0.136	0.122
In education	0.013	0.018	0.018
Missing employment	0.015	0.022+	0.027+
% unemployed	19.524	19.125	21.285
Gross income 10.000 DKK	18.882	18.294+	18.356
Penal conviction 5 years PB	0.044	0.058*	0.052
ADHD medication	-	-	-
Other psychoactive medication	0.061	0.034***	0.062
Observations	1799	2592	597

Source: Own calculations based on data from Statistics Denmark.

Note: Stars indicates significant differences in means between those with consistent treatment patterns and ambiguous/discontinued treatment patterns. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. PB = pre birth of the child. "-" indicates clouded cell because cell size is below 5 observations.

Table A-4. Sensitivity to definition of treatment patterns

	Original	Discontinued treatment		Consistent treatment	
		Max 60 days	Max 120 days	Max 60 day gap	Max 120 day gap
Charged by 18					
Ambiguous treatment	0.024+ (0.014)	0.021 (0.013)	0.030* (0.014)	0.004 (0.024)	0.022* (0.011)
Discontinued treatment	0.053+ (0.028)	0.037 (0.034)	0.070** (0.025)	0.038 (0.034)	0.047+ (0.026)
Number of charges by 18					
Ambiguous treatment	0.009 (0.089)	-0.028 (0.087)	0.062 (0.088)	-0.173 (0.182)	-0.053 (0.075)
Discontinued treatment	0.457* (0.201)	0.231 (0.232)	0.475** (0.183)	0.291 (0.258)	0.419* (0.194)
Observations	7390	7462	7304	7390	7390
Share consistent	0.168	0.186	0.151	0.051	0.335
Share discontinued	0.044	0.024	0.059	0.044	0.044
Convicted by 18					
Ambiguous treatment	0.029+ (0.017)	0.024 (0.017)	0.046** (0.018)	0.002 (0.030)	0.029* (0.014)
Discontinued treatment	0.082* (0.035)	0.045 (0.044)	0.125*** (0.032)	0.062 (0.043)	0.077* (0.033)
Number of convictions by 18					
Ambiguous treatment	0.046 (0.037)	0.02 (0.037)	0.074+ (0.038)	0.008 (0.065)	0.052+ (0.029)
Discontinued treatment	0.233** (0.090)	0.051 (0.094)	0.251** (0.080)	0.204+ (0.105)	0.229** (0.086)
Observations	4332	4374	4275	4332	4332
Share consistent	0.168	0.187	0.150	0.053	0.325
Share discontinued	0.048	0.026	0.064	0.048	0.048
Treatment covariates	X	X	X	X	X
Child demographics	X	X	X	X	X
Other psychoactive medication	X	X	X	X	X
Parental demographics	X	X	X	X	X
Parental SES	X	X	X	X	X
Parental psychoactive medication	X	X	X	X	X

Source: Own calculations based on data from Statistics Denmark.

Note: Table shows main estimates under varying specifications of treatment patterns. In the original specification, discontinued treatment is defined as having filled prescriptions in a period of max. 90 days, and consistent treatment is defined as having no more than 90 days between prescriptions by the age of 15. .

Standard errors are robust. +p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table A-5. Sensitivity to restriction on number of prescriptions			
	Min. 3 prescriptions	Min. 2 prescriptions	Min. 1 prescription
Charged by 18			
Ambiguous treatment	0.024+ (0.014)	0.024+ (0.014)	0.025+ (0.014)
Discontinued treatment	0.053+ (0.028)	0.020 (0.021)	0.003 (0.018)
Number of charges by 18			
Ambiguous treatment	0.009 (0.089)	0.014 (0.088)	0.032 (0.088)
Discontinued treatment	0.457* (0.201)	0.157 (0.139)	0.103 (0.121)
Observations	7390	7756	8218
Share consistent	0.168	0.160	0.151
Share discontinued	0.044	0.081	0.132
Convicted by 18			
Ambiguous treatment	0.029+ (0.017)	0.028 (0.017)	0.030+ (0.017)
Discontinued treatment	0.082* (0.035)	0.033 (0.026)	0.026 (0.022)
Number of convictions by 18			
Ambiguous treatment	0.046 (0.037)	0.049 (0.037)	0.059 (0.037)
Discontinued treatment	0.233** (0.090)	0.077 (0.059)	0.080 (0.050)
Observations	4332	4600	4922
Share consistent	0.168	0.158	0.148
Share discontinued	0.048	0.092	0.151
Treatment covariates	X	X	X
Child demographics	X	X	X
Other psychoactive medication	X	X	X
Parental demographics	X	X	X
Parental SES	X	X	X
Parental psychoactive med.	X	X	X

Source: Own calculations based on data from Statistics Denmark.

Note: Table shows main estimates under varying restrictions on the minimal number of prescriptions purchased. Original specification requires ≥ 3 prescriptions purchased before age 15. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-6. Sensitivity to Heckman corrections

	Heckman correction		
	Original	Two-Step	Maximum-Likelihood
Number of charges by 18			
Ambiguous treatment	0.009 (0.089)	0.004 (0.082)	0.006 (0.089)
Discontinued treatment	0.457* (0.201)	0.462** (0.158)	0.460* (0.200)
Observations	7390	715742	715742
N selected		7390	7390
Lambda		-0.575 (0.240)	-0.342 (0.095)
Rho		-0.224	-0.135
Number of convictions by 18			
Ambiguous treatment	0.046 (0.037)	0.045 (0.038)	0.045 (0.037)
Discontinued treatment	0.233** (0.090)	0.236*** (0.072)	0.236** (0.089)
Observations	4332	575659	575659
N selected		4332	4332
Lambda		-0.176 (0.109)	-0.163 (0.066)
Rho		-0.192	-0.179
All controls	X	X	X

Source: Own calculations based on data from Statistics Denmark.

Note: Table shows main estimates with and without Heckman corrections for selection into treatment. Dummies for municipality at age 6 are used as exclusion restriction, and are significant in predicting selection into treatment, but not in the substantive equation predicting crime-outcomes conditional on other covariates. +p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table A-7. Bounding exercise, charges

	R-squared multiplier = 1.3					
	delta=1	delta=1.5	delta=2	delta=-1	delta=-1.5	delta=-2
Charged by age 18						
Discontinued						
Bias-adjusted beta	0.058	0.060	0.063	0.048	0.045	0.043
CI 95%, ll	0.006	0.008	0.011	-0.011	-0.004	-0.008
CI 95%, ul	0.110	0.112	0.114	0.106	0.094	0.093
Ambiguous						
Bias-adjusted beta	0.034	0.039	0.044	0.014	0.010	0.005
CI 95%, ll	0.006	0.009	0.019	-0.012	-0.015	-0.02
CI 95%, ul	0.061	0.068	0.069	0.041	0.035	0.031
Observations	7390	7390	7390	7390	7390	7390
R-squared	0.084	0.084	0.084	0.084	0.084	0.084
# of charges by age 18						
Discontinued						
Bias-adjusted beta	0.472	0.480	0.488	0.442	0.434	0.427
CI 95%, ll	0.146	0.130	0.162	0.134	0.026	0.085
CI 95%, ul	0.798	0.830	0.814	0.750	0.843	0.768
Ambiguous						
Bias-adjusted beta	0.034	0.047	0.060	-0.014	-0.025	-0.036
CI 95%, ll	-0.136	-0.125	-0.122	-0.184	-0.202	-0.226
CI 95%, ul	0.204	0.219	0.243	0.156	0.152	0.154
Observations	7390	7390	7390	7390	7390	7390
R-squared	0.060	0.060	0.060	0.060	0.060	0.060

Source: Own calculations based on data from Statistics Denmark.

Note: We vary the selection ratio (delta, i.e., the selection on unobservables relative to observables) from 2 (indicating that selection on unobservables is twice as high as selection on observables) to -2 (indicating that selection on unobservables is twice as high and goes in the opposite direction than selection on observables) and calculate the bias-adjusted coefficients using *psacalc* in Stata. R-squared-max is set at 1.3 (as suggested by Oster, 2019) times the R-squared from the regression including all the controls. 95% confidence intervals for bias-adjusted estimates are obtained from bootstrapping (100 replications).

Table A-8. Bounding exercise, convictions

	R-squared multiplier = 1.3					
	delta=1	delta=1.5	delta=2	delta=-1	delta=-1.5	delta=-2
Convicted by age 18						
Discontinued						
Bias-adjusted beta	0.085	0.087	0.088	0.079	0.078	0.076
CI 95%, ll	0.030	0.016	0.010	0.017	0.009	0.002
CI 95%, ul	0.141	0.157	0.166	0.141	0.147	0.151
Ambiguous						
Bias-adjusted beta	0.036	0.039	0.043	0.023	0.020	0.017
CI 95%, ll	0.002	0.004	0.005	-0.009	-0.011	-0.024
CI 95%, ul	0.070	0.075	0.081	0.054	0.050	0.057
Observations	4332	4332	4332	4332	4332	4332
R-squared	0.070	0.070	0.070	0.070	0.070	0.070
# of convictions by age 18						
Discontinued						
Bias-adjusted beta	0.239	0.242	0.245	0.227	0.224	0.221
CI 95%, ll	0.082	0.059	0.067	0.032	0.072	0.053
CI 95%, ul	0.396	0.425	0.423	0.422	0.377	0.39
Ambiguous						
Bias-adjusted beta	0.059	0.065	0.072	0.034	0.029	0.023
CI 95%, ll	-0.016	-0.006	0.004	-0.043	-0.046	-0.044
CI 95%, ul	0.133	0.137	0.141	0.112	0.103	0.09
Observations	4332	4332	4332	4332	4332	4332
R-squared	0.075	0.075	0.075	0.075	0.075	0.075

Source: Own calculations based on data from Statistics Denmark.

Note: We vary the selection ratio (delta, i.e., the selection on unobservables relative to observables) from 2 (indicating that selection on unobservables is twice as high as selection on observables) to -2 (indicating that selection on unobservables is twice as high and goes in the opposite direction than selection on observables) and calculate the bias-adjusted coefficients using *psacalc* in Stata. R-squared-max is set at 1.3 (as suggested by Oster, 2019) times the R-squared from the regression including all the controls. 95% confidence intervals for bias-adjusted estimates are obtained from bootstrapping (100 replications).

Figure A-1. Coefficients for different sample restrictions

a) Charged by age 18

b) # charges by age 18

c) Convicted by age 18

d) # convictions by age 18

Source: Own calculations based on data from Statistics Denmark.

Note: Figure shows sensitivity to different sample restrictions. 1) dropping those who start treatment before age 5, 2) dropping those who start treatment less than one year prior to outcome is measured/treatment patterns are defined, 3) dropping those born prior to 1990, 4) dropping females, 5) dropping those with immigrant background (immigrant or descendant). Lines indicate 95% confidence intervals. N=7390 for charge outcomes, N=4332 for conviction outcomes in the full samples.