Effects of ADHD Treatment Sequencing on Patient Health and Social Outcomes

Anna Chorniy
Clemson University *†

PRELIMINARY DRAFT
WORK IN PROGRESS

September 2014

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a common mental chronic condition that negatively affects noncognitive abilities. This paper estimates short- and long-term effects of various treatment strategies on human capital accumulation in children and adolescents with ADHD.

The medical literature suggests that ADHD can seldom be cured, but in most cases, its symptoms can be effectively managed. In order to relieve symptoms of the condition and augment the stock of noncognitive skills, patients take ADHD medications and/or attend psychotherapy sessions. While on treatment, they can learn planning and self-control skills that help them better manage their condition in the future.

Poor noncognitive abilities may lead to such negative health and social outcomes as teen pregnancy, contraction of STDs, injuries, and the onset of depression. Accumulation of ADHD "management" skills reduces the probability of adverse events in the future. Hence, there is a link between previous treatment and future health outcomes, as well as the future treatment choices.

*Author correspondence: achorni@clemson.edu

†I owe special thanks to Tom Mroz, Daniel Miller, and Michael Maloney for their guidance. I am also grateful to Donna Gilleskie for comments on previous drafts and for sharing her Fortran code with me. Further, I thank the participants at the Clemson Industrial Organization seminar and Labor Economics seminar for helpful suggestions. Finally, I would like to thank a number of people who helped me obtain access to the data: Mark Harouff, Heather Kirby, and Joe Magagnoli.

Using SC Medicaid claims data for 2003-2012, I model dynamic treatment decisions for children with ADHD and subsequent adverse health and behavioral outcomes. I use the estimates to simulate the cost of treatment to Medicaid under various treatment sequences, including the cost of poor adherence and late diagnosis.

1 Introduction

One of the important insights from the literature on child development is that gaps in abilities that form early in life persist and explain a large array of differentials in adult outcomes. Conti and Heckman (2014) provide an extensive review of the empirical evidence on the effects of the two dimensions of child well-being, cognitive and noncognitive skills on educational attainment, asocial and risky behaviors, and health. These recent studies emphasize the importance of modeling of multidimensional capabilities as opposed to earlier literature on human capital development that concentrated on child cognitive abilities, often measured by IQ to explain the outcomes later in life.

One of the earliest studies to account for the latent noncognitive skills is by Heckman et al. (2006). They find that both cognitive and noncognitive abilities affect wages, schooling, work experience, occupational choice, and participation in a range of adolescent risky behaviors. These results have important policy implications. Most interventions target children's IQ but not their noncognitive abilities. The Perry Preschool experiment that improved a great variety of outcomes did not result in the improvements in IQ scores. Heckman et al. (2006) argue that the improvement came from social skills.

The focus of this paper is the investment in noncognitive skills of disadvantaged children, who were born with the attention-deficit/hyperactivity disorder (ADHD). ADHD is a chronic mental condition that is common among children and adolescents. There are three types of ADHD: predominantly hyperactive, inattentive, or a combination type. Patients of the hyperactive type lack self-control, patience, and demonstrate immature behavior that is inconsistent with their age group. Inattentive type patients have a poor ability to concentrate their attention, complete tasks, and are forgetful. In other words, ADHD is an impairment of noncognitive skills.

In order to reduce the gap in abilities of children with ADHD compared to their non-ADHD peers, the condition should be managed. Once a child is diagnosed, her family can invest in medical treatment. While on treatment, the child is able to improve her outcomes in the short-run and also accumulate cognitive and noncognitive skills and improve the long-run outcomes. In particular, I am interested in the effectiveness of medical treatment and its importance early in life versus later in life, its consistency, and specific treatment choice sequence for development of noncognitive abilities. Successful ADHD treatment reduces the probability of adverse outcomes in the future.

The most recent 2011/12 National Survey of Children's Health reports that over 5 million children (7.9%) aged 2–17 are currently diagnosed with ADHD in the U.S. Over 68% of these children are taking medications for the disorder.[1] However, very little is known about the relative effectiveness of available treatments and about their overall effects on health, behavioral, and school outcomes, especially in the long-run.

In recent years the media launched an attack on the rapidly rising trend of ADHD diagnoses and prescriptions. In an avalanche of articles mental health professionals argue that ADHD drugs are overprescribed. The drugs are said to be overused by students due to their immediate effect on the ability to concentrate. However, they warn the reader that the life-long consequences of taking these medicines are unknown. Moreover, there is no evidence of any long-term positive effects on educational or behavioral outcomes, but there is a number of worrisome side effects as, for example slowdown in growth and addiction.[2]

This stance on ADHD medications comes from the fact that clinical studies are predominantly short, lasting just a few weeks. They do not have enough data to reach any serious conclusions. Excluding the extreme cases when ADHD drugs are taken solely on the test, the argument misses the dynamic nature of the human capital development. Even if the effects of treatment are short-lived, the child has an opportunity to learn how to manage her condition while on treatment and accumulate social skills that will improve her outcomes later in life.

There is an emerging health economic literature on the effects of ADHD treatment on short- and long-term outcomes. For example, Currie et al. (2014) use a quasi-natural experiment that lowered prices on all prescription drugs in Quebec, Canada but not in other provinces. They find little evidence of positive effects on academic outcomes and even some evidence of negative impact of treatment on grade repetition and math scores, and emotional stability of girls. Dalsgaard et al. (2014) look at health services utilization (hospital visits) and behavioral outcomes (crime), using the variation in the doctor propensity to prescribe pharmacological treatment as an IV. They find a positive effect of treatment on patient health and behavior. Treated children had fewer hospital visits, driven by fewer injuries,

[1] National Survey of Children's Health. NSCH 2011/12. Data query from the Child and Adolescent Health Measurement Initiative, Data Resource Center for Child and Adolescent Health website. Retrieved 09/25/2014 from www.childhealthdata.org.

[2] See for example, "Ritalin Gone Wrong." by Sroufe, L. Alan. The New York Times, January 28, 2012; "Risky Rise of the Good-Grade Pill." Schwarz, Alan. The New York Times, June 9, 2012; "Drowned in a Stream of Prescriptions." Schwarz, Alan. The New York Times, February 2, 2013; "A Nation of Kids on Speed." Cohen, Pieter and Rasmussen, Nicholas. The Wall Street Journal, June 16, 2013; and "The Truth About Smart Drugs" by Marek Kohn, BBC, July 29, 2014.

and they also had fewer encounters with the police. Using the same IV applied to a sample of children and young adults enrolled in SC Medicaid in 2003–2012, Chorniy and Kitashima (2014) find that ADHD treatment reduces the probability of teenage pregnancy, another adverse behavioral outcome relevant for ADHD population.

I contribute to this literature by explicitly capturing the dynamic nature of the problem. I concentrate on the long-run effects of ADHD treatment on the consequences of the risky sexual behavior (teen pregnancy and STDs) and on health outcomes (injuries and early onset of depression). I also take a more precise approach to treatment by distinguishing between meaningfully different pharmacological treatments as well as by including behavioral therapy into the choice set. This approach allows me to test for the importance of a particular sequence of treatments.

The problem of a choice of treatment is also addressed in the recent literature on choice under uncertainty (see Crawford and Shum (2005), Dickstein (2014), and Saxell (2013)). They use learning models with Bayesian updating to model the process of patient search for most suitable drug. This process relies only on the choice made in the previous period. Instead, I focus on the entire treatment sequence. This approach allows for testing a hypothesis that some treatments are more valuable at the start period of treatment and others are more suitable for an established patient. To my knowledge, this particular approach has not been used before.

These dynamic studies also suffered from the lack of data on patient outcomes. They typically assume that a patient is cured when she exits treatment. My data allow me to introduce more realistic measures of treatment effectiveness – a number of behavioral and health outcomes that I identified from the medical literature (see **?** for a detailed review) and literature on child well-being (e.g. Heckman et al. (2006)). The four adverse outcomes are identifiable in my data. They are outcomes associated with risky sexual behavior (teenage pregnancy and STDs), health outcomes associated with poor attention and hyperactivity (injuries), and an onset of depression.

I use a large panel dataset of SC Medicaid claims in 2003–2012 that is enriched by data from the patients' birth certificates. According to the 2011/12 National Survey of Children's Health, in SC over 10% of children between 2 and 17 years old are diagnosed with ADHD and 63% of them are on medical treatment for their condition. In the Medicaid population, these statistics are even higher. For example, in a cohort of children born in 1996 and eligible for SC Medicaid between 2003–2012 over 23% of children between 3 and 19 years old have been diagnosed with ADHD during the sample period. Moreover, about 80% of those diagnosed with the condition were prescribed pharmacological or non-pharmacological treatment.

Having a long panel allows me to use an empirical approach that is commonly used for dynamic processes that have unobserved heterogeneity (see Mroz and Savage (2006) and Yang et al. (2009)). I model and estimate simultaneously the equations for the event of the initial ADHD diagnosis, for treatment decision, and adverse events. The discrete factor random effects estimator is beneficial in this setting because it can be used to control for endogeneity biases in nonlinear models where fixed effect estimators would be inconsistent. For comparison, I estimate all the events and outcomes as single equations.

To summarize, I am looking at ADHD treatment in the context of the model of investment in child development. I am particularly interested in the long-run effects of treatment on adverse health and behavioral outcomes; relative importance of early investments in child development and their consistency.

2 Conceptual framework

2.1 ADHD and noncognitive ability

Every child is born with a multidimensional endowment of abilities. They include cognitive (e.g. IQ, memory) and noncognitive (e.g. self-control, patience, time preference) skills (cite). Most recent medical research suggests that genetic and neurological factors are the greatest contributors to the ADHD (see Barkley (2006) for an extensive review). Due to their condition, children who suffer from ADHD have a relatively low initial stock of noncognitive skills.

Poor noncongitive abilities may lead to a number of negative health and social outcomes, such as teen pregnancy, contraction of STDs, injuries, and the onset of depression (as described in Section 2). The medical literature suggests that ADHD can seldom be cured, but in most cases, its symptoms can be effectively managed. In order to relieve symptoms of the condition and augment the stock of noncognitive skills, patients take ADHD drugs and/or attend psychotherapy sessions. The pharmacology of ADHD medicines is such that the effect of treatment goes away as soon as the patient stops taking them. However while on treatment, pharmacological or behavioral, patients are able to accumulate human capital. They can learn planning and self-control skills in order to better manage their ADHD symptoms in the future. Accumulation of ADHD "management" skills reduces the probability of adverse events in the future. This makes up a link between previous ADHD treatment, current stock of noncognitive skills, and future health and social outcomes, as well as the future treatment choices.

ADHD treatments can only be prescribed after the initial diagnosis. According to the

American Academy of Pediatrics guidelines[3], primary care clinicians should evaluate children between 4 and 18 years old for ADHD if they show some of the symptoms.[4] Since ADHD is a hereditary rather than acquired condition, the timing of diagnosis depends on the severity of symptoms. Also, it seems that hyperactive types are more likely to be diagnosed earlier than inattentive types simply because inattentiveness might be confused with poor cognitive skills but if a child is acting up, parents and teachers are more likely to do something about it (cite).

Earlier rather than later in life diagnosis might be beneficial if Heckman theory is right. However, they say noncognitive skills are easier to improve later in life than cognitive skills. Even if a child did not experience any adverse health events early in life, the fact that ADHD has not been controlled till her teens will increase the likelihood of negative outcomes in adolescence.

To formalize the model, I use the general theoretical framework of the technology of skill formation and investment in human capital laid out in Cunha and Heckman (2007) and Cunha et al. (2010). I will concentrate on noncognitive skills only and refer to ADHD treatments as investment into these skills.

2.2 Timing

The model timeline can be divided into three parts. First, when a child is born she receives an initial endowment of noncognitive skills, θ_0, that depends on genetic and environmental factors. Second, once a child reaches the age of 3, she can be tested for and diagnosed with ADHD (D_t) if she has some of the symptoms. Finally, once the patient who has the condition, is diagnosed, she can be prescribed a medical treatment (T_t) to relieve symptoms of ADHD. The treatment augments the stock of noncognitive skills that feeds into the next time period. It also affects the probability of the adverse events that may be realized in the next period (E_{t+1}).

Figure 1 depicts the dynamics of the stock of noncognitive skills and the probability of adverse events linked to poor noncognitive abilities. It shows a representative year of the time period, when a patient had been diagnosed with ADHD. At the start of the year t this patient has information on her stock of noncognitive skills, adverse events that were realized in the past years, and past treatment (if any). During year t she will be making decisions

[3]Subcommittee on Attention-deficit/hyperactivity disorder, steering committee on quality improvement and management, "ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents", *Pediatrics*, 2011.

[4]Since a number of ADHD prescription drugs are approved by the FDA for use in children as young as 3 years old (e.g. Adderall, Adderall XR), I will use age 3 as the first time period when a diagnosis can be made and treatment initiated.

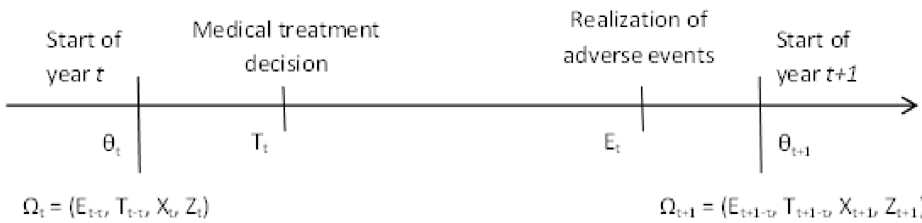

Figure 1: Decision Timeline

on treatment for her condition. Also, adverse events will or will not be realized. Then, by year $t+1$ the individual will have an updated stock of noncognitive skills due to medical treatment.

In sum, at the beginning of each time period (year) an individual has the following information that influences her treatment choice in that time period: current stock of noncognitive skills (θ_t), occurrence of adverse events in the past years ($E_{t-\tau}...E_{t-1}$), ADHD diagnosis status (D_{t-1}) and if diagnosed, what treatment have they undergone in the previous periods ($T_{t-\tau}...T_{t-1}$). They also have information on exogenous supply side variables, such as treatment prices ($\vec{p}$) and characteristics; physician characteristics, and a number of individual and time-specific variables (location, family size and composition). Finally, they have information on current and lagged unobservable to the researcher variables that feed into the optimization problem.

Prior to the initial diagnosis of ADHD no treatment can be prescribed. The timeline for a representative year before the initial diagnosis is similar to the one shown on Figure 1, except for there is no treatment decision to make. Individuals transition into the next period with an unaffected stock of noncognitive skills ($\theta_{t+1} = \theta_t$).

In what follows I describe an empirical specification for every component of this dynamic system.

3 Empirical specification

In this paper, I focus on treatment of ADHD as a way of investing in the child's noncognitive skills. The model is characterized by a decision on medical treatment, four health and behavioral outcomes (injuries, teen pregnancy, STDs, and onset of depression), an event of initial diagnosis with ADHD, and an initial condition for the stock of noncognitive skills. Each of the decisions, events, outcomes, and intial conditions are specified with an equation. In this section I detail them all.

3.1 Initial Condition

Children are born with an initial endowment of health and abilities that include cognitive skills (θ_0^C) and noncognitive skills (θ_0^N). Initial endowments of skills are influenced by family environments and genetic factors (cite Olds'02 and Levitt'03, see CHS10).

$$\theta_0 = (\theta_0^C, \theta_0^N) \tag{1}$$

Although cognitive and noncognitive skills are interrelated, due to the data limitations I will focus solely on noncognitive skills accumulation.[5] There is some evidence that this will not bias my results (cite).

I specify the initial endowment of noncognitive skills as depending on the observed parental characteristics at birth: mother's age (a_i), race (r_i), education level (s_i), history of mental disorders (h_i), and unobserved characteristics (ε_i).

$$\theta_{i0} = \xi_1^m a_i^m + \xi_2^m r_i^m + \xi_3^m s_i^m + \xi_4^m h_i^m + \varepsilon_i, \tag{2}$$

where i indexes individuals with ADHD and m superscript corresponds to the child's mother. I do not directly estimate Equation (2). Instead, it feeds into the equations on the timing of the first diagnosis and outcome equations.

3.2 Treatment

The initial stock of noncognitive skills can be altered with investments in child's development. Higher noncognitive abilities reduce the likelihood of adverse health and social outcomes that children and adolescents with ADHD are prone to. One way to improve noncognitive abilities of children with ADHD is to treat their disabling mental condition. In each period when a treatment is administered, ADHD symptoms subside and the child's ability to concentrate and control her impulses improves. Additionally, while on treatment, the child learns how to manage ADHD symptoms better. In other words, ADHD treatment alters the stock of noncognitive skills in current and future periods.

In the model, pharmacological and behavioral ADHD treatments are assumed to be the only investments available to parents to improve their child's noncognitive skills. I assume that both physician and parents are perfect agents of the child.[6]

[5] One of the limitations of claims data is the lack of such data as cognitive abilities or other parental investments in child development.

[6] Although there is emerging evidence that parents are important agent here and it is important to account for redistribution of wealth in families (see for example xxx), one drawback of claims data surfaces. I can't look at it. Doctor agency was also shown to be important in choosing medical treatments (see Dickstein, etc) but I leave it for the future work.

There is no depreciation in the model. If a child is not treated in the current period, her stock of noncognitive skills in the next period will remain on the same level as it was in the earlier period. According to the medical literature, ADHD can seldom be cured (see for example, xxx). However, children can learn how to manage their condition, when they reach a high enough level of noncognitive skills.

At the beginning of year t for patient i diagnosed with ADHD has the stock of noncognitive skills, θ_{it}. It depends on the stock of abilities at the beginning of previous year, $t-1$ and whether the child was treated, M_{it-1}.

$$\theta_{it} = f(\theta_{it-1}, M_{it-1}) \tag{3}$$

Alternatively, it can be written as follows.

$$\theta_{it} = f(\theta_{i0}, M_{it-1}, M_{it-2}, \ldots, M_{it-9}) \tag{4}$$

where t is a survey year, $t = 1, \ldots, 10$.

include branded status, side effects, dosing frequency, all with random coefficients?

ADHD treatment decision depends on the lifetime value of improving noncognitive skills in this period. It includes contemporaneous utility and the expected present discounted value of the future utility conditional on the choice of treatment in the current period. The current indirect utility depends on the severity of ADHD (θ_{it}), adverse events realized in the past (three) years ($E_{t-3}, ..., E_t$), prices of available treatments ($\vec{p_t}$), other drug characteristics (d_t), provider characteristics (b_{jt}), location (l_{it}), and person/period-specific variables, z_{it}. Thus, the value function has the following arguments:

$$V_T(\theta_{it}, E_{t-3}, ..., E_t, \vec{p_t}, b_{jt}, z_{it}|D_{it}) \tag{5}$$

$$V_T(\theta_{i0}, T_{t-1}, E_{t-3}, ..., E_t, \vec{p_t}, b_{jt}, z_{it}|D_{it}) \tag{6}$$

The choice of treatment after the diagnosis, T_{it} depends on

$$T_{it} = \omega_T \theta_{it} + \beta_T l_{it} + \gamma_T z'_{it} + \sum_{\tau=1}^{3} \delta_E E_{it-\tau} + \kappa_T b_{it} + \lambda \vec{p_{T,t}} + \epsilon_{T,it} \tag{7}$$

Substituting θ_{it} from the equation 3, I get the following expression:

$$T_{it} = \alpha_{T1}a_i^m + \alpha_{T2}r_i^m + \alpha_{T3}s_i^m + \alpha_{T4}h_i^m + \beta_T l_{it} + \gamma_T z'_{it} + \sum_{\tau=1}^{3} \delta_T E_{it-\tau} + \kappa_T b_{it} + \lambda \vec{p_{T,t}} + \sum_{\tau=1}^{9} \delta_T T_{it-\tau} + \epsilon_{T,it} + \varepsilon_i \quad (8)$$

In this model, the present and future are linked through the process of human capital accumulation. Parents invest in medical treatment of their children with ADHD.

My model allows to test a hypothesis of the importance of early investments into child's development that is an important finding in the recent literature on child development (e.g. Cunha and Heckman (2007)) as well as in medical literature (see for example, xxx). Cunha and Heckman (2007) and Cunha et al. (2010) state that investments in child health are productive at all time periods, but their productivity differs. Treatments early in life are likely to be more productive than later in life.

3.3 Adverse events

Low level of noncognitive skills is an important determinant of poor educational, labor market outcomes, and social outcomes (Heckman et al. (2006)). Following medical (cite, cite) and economic literature on child development (**?**, Heckman et al. (2006)), I concentrate on the following adverse events that are typical for children and young adults with ADHD: teen pregnancy, contraction of STDs, injuries, and an early onset of depression. The outcomes of risky sexual behavior are age-specific. I only model this outcome for the female individuals older than 11 and younger than 17(19) years of age. The outcome on the onset of depression becomes relevant at age xx. In a companion paper (**?**) we look at the educational and behavioral outcomes of treatment.

3.3.1 Risky sexual behavior

Adolescents with untreated ADHD struggle to control their impulses and plan ahead. Moreover, these teens often struggle with low self-esteem, often seek affirmation through the sexual attentions of boys in an effort to compensate for feelings of inadequacy in other areas of their life (Arnold (1996)). [7] Their condition makes them more likely to become sexually active earlier than their peers, not use or inconsistently use birth control, and most importantly, have more partners on average than unaffected peers have (Kessler et al. (1997),blo

[7] adolescent girls' symptoms of ADHD often worsen due to the hormonal changes at puberty (Resnick (2005)).

(Accessed on May 16^{th}, 2014)). The two adverse events associated with risky sexual behavior are teen pregnancy and contraction of a sexually transmitted disease (STD).

In the U.S. in 2013 274,641 babies were born to mothers aged 15–19 years and 3,108 babies to mothers under 15 years old, a live birth rate of 26.6 and 0.3 per 1,000 women in these age groups (cdc may 2014).[8] About 80% of teenage births are unplanned or unwanted (link1) and only 59% of them ended with a live birth in 2008 (Finer and Zolna (2011)).

Teenage pregnancy is a significant negative social and health outcome. Adolescent mothers are more likely to be single (cite), be on welfare (cite) and have a hard time getting off welfare (Kessler, #9.). Teenage pregnancy is also associated with negative consequences for the mother later in life (low educational attainment, poor employment outcomes, and marital instability (see 7,8,11) and poor child outcomes (low birth weight, delay in cognitive development, school problems, behavioral disorders, and becoming teenage parents themselves, see 2,3,4).

In contrast to the trend in teen pregnancy, a declining trend of STD cases in the U.S. was reversed in early 2000s. In 2012, there have been 49,903 cases of STDs (16.0 per 100,000 population). Adolescents ages 15-24 account for nearly half of the new cases of STDs each year (Sheet).

ADHD treatment reduces symptoms of the condition and should be able to reduce the probability of these adverse outcomes. However, for it to be most effective it should be managed with care because of reproductive issues and the potential impact of menstrual cycles on drug effectiveness (Resnick (2005)).

3.3.2 Injuries

Inattentiveness, difficulty in assessing potential outcomes, and motor incoordination are also a frequent cause of accidental injuries (e.g. fractures) for patients with ADHD (Barkley (2006)). Over half of them is estimated to be accident-prone with their injuries being not only more severe but also more frequent (Barkley (2006), Swensen et al. (2004).

In particular, researchers find that ADHD adolescents are more likely to have at least one car crash and they are more severe than for their non-ADHD peers. They are also more often at fault of such accidents (Barkley (2006), Weiss and Hechtman (1993)).

In their work on long-term consequences of ADHD treatment, **?** find that pharmaceutical treatment of ADHD results in fewer hospital visits.

[8] CDC calls births between 10 and 19 teenage births. It is divided into categories 10-14, 15-17, and 18-19.

3.3.3 Onset of depression

[in progress]

3.3.4 Empirical specification

I am interested in the effect of treatment on these outcomes. I specify an equation for an occurrence of each outcome. These events are modeled as discrete outcomes. are commonly associated with a variety of negative health and behavioral outcomes, E_{it}.

$$E_{it} = \omega_E \theta_{it} + \beta_E l_{it} + \gamma_E z'_{it} + \epsilon_{E,it} \tag{9}$$

Substituting θ_{it} from the equation 3, I get the following expression:

$$E_{it} = \alpha_{E1} a_i^m + \alpha_{E2} r_i^m + \alpha_{E3} s_i^m + \alpha_{E4} h_i^m + \beta_E l_{it} + \sum_{\tau=1}^{9} \delta_E T_{it-\tau} + \epsilon_{E,it} + \varepsilon_i \tag{10}$$

3.4 First diagnosis of ADHD

An eligible child-enrollee can be tested and diagnosed with ADHD at a medical provider office. Any doctor is able to diagnose and prescribe treatments (except for physcologists[double-check!]). In order to be diagnosed ($D^*_{it} = 1$), the test should reveal at least six of the inattention symptoms and/or at least six of hyperactivity-impulsivity symptoms that "have persisted for a least 6 months to a degree that is maladaptive and inconsistent with developmental level."[9] It is extremely rare for a child to be diagnosed before age 3 because the symptoms are not apparent at this age.

Whether ADHD is diagnosed in any given year depends on the contemporaneous stock of noncognitive skills (θ_{i0} at age 3 and θ_{it} at age 4, 5,..., 21) and on the history of adverse events ($E_{it-\tau}$). Since at age 3 only three periods of history are available, I will use three lags in my default specification. I will also test specifications that extend further into the past.

[9] The American Psychiatric Association publishes the Diagnostic and Statistical Manual of Mental Disorders (DSM), where it sets criteria for the classification of mental disorders. It is the standard classification of mental disorders used by mental health professionals in the United States. The DSM consists of three major components: the diagnostic classification, the diagnostic criteria sets, and the descriptive text. The most current version is DSM-5 published in May 2013, a revision of DSM-IV-TR that came out in 2000.

Besides noncognitive skills level and history of adverse events associated with ADHD, the probability of being diagnosed depends on individual, z'_{it} (e.g. age, race, gender), family, f'_{it} (e.g. number of adults and children in the family, county), and medical provider, b'_{it} (e.g. location, specialty) characteristics.

Diagnosis is specified as a latent variable and can be written as follows.

$$D^*_{it} = \omega_D \theta_{i0} + \sum_{\tau=1}^{3} \delta_D E_{it-\tau} + \beta_D z'_{it} + \beta_D f'_{it} + \gamma_D b'_{it} + \epsilon_{D,it} \tag{11}$$

where $D_{it} = 1$ if $D^*_{it} > 0$ and 0 otherwise, t indexes the year of diagnosis, $t = 1, 2, ..., 10$ that correspond to the survey period 2003-2012.

Substituting θ_{i0} from the equation 3, I get the following expression:

$$D^*_{it} = \alpha_{D1} a^m_i + \alpha_{D2} r^m_i + \alpha_{D3} s^m_i + \alpha_{D4} h^m_i + \sum_{\tau=1}^{3} \delta_D E_{it-\tau} + \beta_D z'_{it} + \beta_D f'_{it} + \gamma_D b'_{it} + \epsilon_{D,it} + \varepsilon_i \tag{12}$$

3.5 Likelihood function

Following Mroz and Savage (2006) and Yang et al. (2009) I use the DFML method to control for heterogeneity and endogeneity by integrating out the unobserved factors μ_i and η_{it}. the contribution to the likelihood of the individual i in year y is:

$$\begin{aligned} L_{it}(\Omega|\mu_i, \eta_{it}) = & \left[\Pr\{D_{it} = 1|\mu_i, \eta_{it}\} \cdot \Pr\{T_{it} = 1|\mu_i, \eta_{it}\}^T_{it} \cdot \Pr\{T_{it} = 0|\mu_i, \eta_{it}\}^{(1-T_{it})}\right]^{D_{it}} \cdot \\ & [\Pr\{D_{it} = 0|\mu_i, \eta_{it}\}]^{(1-D_{it})} \cdot \\ & [\Pr\{E_{it} = 1|\mu_i, \eta_{it}\}]^{E_{it}} \cdot [\Pr\{E_{it} = 0|\mu_i, \eta_{it}\}]^{(1-E_{it})} \end{aligned} \tag{13}$$

where Ω is a vector of parameters to be estimated.

Approximating the continuous distributions of η_i and ν_{it} with mass points η_{1j}, $j = 1, ..., J$, η_{2k}, $k = 1, ..., K$, and vector ν_m, $m = 1, ..., M$, the unconditional contribution to the likelihood function of individual i is:

$$L_i(\Omega, \Gamma) = \sum_{j=1}^{J} p_{1j} \sum_{k=1}^{K} p_{1j} \cdot f_{ins}(ins_i|\eta_{1j}, \eta_{2k}) \cdot \prod_{t=1}^{10} \sum_{m=1}^{M} p_{3m} L_{it}(\Omega|\mu_{1j}, \mu_{2k}, \eta_m) \tag{14}$$

where $p_{gr} = \{\nu_{gi} = \mu_{gr}\}$ for $\mu_{gr} \in R$ and $g = 1, 2$,

$p_{3m} = \Pr\{\eta_i = \eta_m\}$ for $\eta_m \in R^8$, where $f_{ins}(.)$ is the density for the initial condition that describes the level of noncognitive skills at birth, and Γ is the vector containing the parameters of the discrete distributions.

I use Fortran programs to obtain maximum likelihood estimates.

For comparison purposes, I first estimate single-equation specifications for every outcome of interest, as well as for the event of the first ADHD diagnosis and decision on medical treatment. Currently, all my dependent variables are specified as discrete, so I use probit regressions. I will compare the results from these single-equation regressions to the results obtained using the discrete factor maximum likelihood approach.

3.6 Identification

In progress.

4 Data

4.1 Medicaid Claims

I use a large panel dataset of medical claims filed to and paid by SC Medicaid in 2003-2012. The Medicaid program in SC is a one of the major health insurance providers with about 20% of the state population being active enrollees and over $5 billion in spending (source, 2009 data). Medicaid is a means-tested program that target population is low-income families, disabled, aged, and blind individuals, and pregnant women.

In SC, the Children's Health Insurance Program (CHIP) is a part of Medicaid. Under the general rules, families with income below 150% of the Federal Poverty Level (FPL) are eligible for Medicaid and children from families with income is below 200% of the FPL are eligible for Medicaid coverage through CHIP. Half of the Medicaid insured are children.

My data include Medicaid monthly eligibility status, hospital, outpatient, and pharmacy claims for individuals who were diagnosed with ADHD between 3 and 21 years old. A hospital claim may have record of up to 9 diagnosis codes and an outpatient claim – up to 3 codes. The Medicaid program uses standard ICD9 codes, where ADHD diagnosis is coded as ICD9 code 314.xx. Every patient in my sample has at least one claim with ADHD diagnosis.

Claim records do not contain information on patient health history. In order to avoid difficulties associated with left-censoring, I take steps to determine the instance of the first ADHD diagnosis. I excluded cases when the first claim with ADHD diagnosis in the data is within half a year from the eligibility date and cases when a stimulant or non-stimulant medication prescription is filled prior to the first documented ADHD-related physician visit.

I also use ICD9 diagnosis codes to find the adverse events in the data. Pregnancies are defined as the first instance when one of the pregnancy-related codes appears (), same of STDs. The ICD9 codes for injuries were borrowed from (cite).

Note that pharmacy claims have no diagnosis record. Instead, I use medical literature to construct a choice set for individuals who have ADHD (see section xxx) and NDC codes[10] to map these choices.

Once eligibility is established, Medicaid coverage is available for an enrollee for a 12-month period (unless the enrollee becomes ineligible during this time), after which the eligibility needs to be reconfirmed. An eligible individual who received services prior to the actual enrollment, can be covered retroactively for up to three months prior to the month when eligibility was established.

In the data, about xx% of the individuals in the sample have lapses in eligibility that are on average xx months long. Since I do not have any information on these individuals when they are ineligible for Medicaid, I only select Medicaid enrollees who are consistently eligible for the program. I disregard lapses in eligibility that last under three months. For inconsistent eligibility periods that result in longer lapses in coverage, I only keep the medical history to the point prior to the lapse. Furthermore, I exclude patients who were covered for less than a year between 2003 and 2012.

SC Medicaid has two components: "traditional" fee-for-service coverage and a number of managed care plans. Due to the differences in report requirements the complete information on all services provided to an enrollee are only available for those who have the fee-for-service coverage only. However, mental health is one of the "carved-out" conditions that are covered by the fee-for-service Medicaid component even if an individual is enrolled into a managed care plan. I perform the estimation on the entire sample of population and do a robustness check by selecting only those individuals who were enrolled into the traditional Medicaid.

In 2013, most eligible individuals faced a small copay per doctor visit ($3.30), per prescription ($3.40 for adults over 19 years old and zero otherwise), per hospital stay ($25). The state maintains preferred drug lists for medicines that do not require prior authorization, all other drugs may be covered if a doctor-filed authorization request is approved. The quantity restrictions are also common with a typical prescription capped at a 30-day supply.

Finally, the claims data is enriched by additional information from enrollees' birth certificates. It includes mom's de-identified ID, age, race, and education level.

Summary statistics on the overall sample is presented in Table 1. Tables 2 and 3 show summary statistics on individual characteristics and health and behavioral outcomes

[10]NDC codes are...

respectively.

Table 1: SC Medicaid Sample of ADHD and non-ADHD enrollees by birth cohort.

Birth cohort	**Eligible for SC Medicaid**	**Diagnosed with ADHD**	**Sample of undiagnosed enrollees**
1978	25,741	0.07%	
1979	27,466	0.17%	
1980	29,458	0.27%	
1981	29,822	0.32%	
1982	30,800	0.55%	
1983	34,543	0.62%	
1984	35,786	1.03%	
1985	37,434	2.13%	
1986	37,440	3.56%	
1987	37,272	5.70%	
1988	37,815	8.01%	
1989	38,654	10.22%	
1990	38,420	12.83%	
1991	37,501	15.27%	
1992	35,992	17.64%	
1993	34,298	19.90%	
1994	33,763	21.14%	
1995	33,113	22.21%	
1996	32,580	23.77%	
1997	33,759	23.26%	
1998	35,615	23.02%	
1999	37,560	21.67%	
2000	40,197	20.50%	
2001	42,510	18.52%	
2002	42,150	17.71%	
2003	41,984	16.69%	
2004	43,111	14.72%	
2005	44,065	11.63%	
2006	48,614	7.76%	
2007	49,782	3.92%	
2008	51,126	1.25%	
2009	49,455	0.20%	
Total	**1,207,826**	**131,007**	**150,000**

Notes: The ADHD sample includes every SC Medicaid enrollee, who was diagnosed with ADHD at any age between 3 and 25 years old in 2003-2012.

Table 2: Summary stats. SC Medicaid, 2003-2012.

	N obs	**Mean**	**Median**	**Std**	**Min**	**Max**
Individual characteristics						
Male	51,767	0.66			0	1
Race: White	25,930	0.52			0	1
Race: Black	22,021	0.44			0	1
Age at 1st ADHD diagnosis	51,767	8.76	8.00	3.81	3	21
Mother's characteristics						
Age when gave birth	28,970	23.27	22.00	5.42	11	46
Newborn's BMI	6,884	28.92	27.46	8.23	13	90
Race: White	14,666	0.51			0	1
Race: Black	13,642	0.47			0	1
Educ: Less than HS	28,847	0.05			0	1
Educ: Some HS	28,847	0.37			0	1
Educ: HS diploma	28,847	0.40			0	1
Educ: Some college	28,847	0.13			0	1
Educ: College	28,847	0.05			0	1
Educ: Grad school	28,847	0.00			0	1
Family characteristics						
N adults, by patient	51,767	1.03	1.00	0.65	0	5
N children, by patient	51,767	1.99	2.00	1.02	0	9

Notes: The ADHD sample includes every SC Medicaid enrollee, who was diagnosed with ADHD at any age between 3 and 25 years old in 2003-2012.

Table 3: Summary stats. SC Medicaid, 2003-2012.

	N obs	Mean	Median	Std	Min	Max
Outcomes: Risky sexual behavior						
Age when became pregnant	2,050	17.94	18.00	2.23	11	25
Teen pregnancies, 11-19	717	16.96	17.00	2.61	11	19
Teen pregnancies, 11-17	1,246	15.87	16.00	1.53	11	17
Age when contracted STD						
Outcomes: Depression						
Age when diagnosed with depression						
Outcomes: Injuries						
N outpatient visits	51,767	0.29	0.10	0.79	0	58
N inpatient stays	51,767	0.19	0.10	0.28	0	9
Inpatient stay days	51,767	0.26	0.10	1.26	0	197
N ADHD-related injuries (total)	51,767	0.03	0.00	0.21	0	34
Inpatient days due to inj+ADHD	51,767	0.01	0.00	0.18	0	23
Medical treatment						
N Rx	51,767	3.01	1.63	3.66	0	32
N Btherapy sessions	51,767	1.40	0.00	4.45	0	181
Years of data	51,767	7.54	8.00	2.62	1	10

Notes: The ADHD sample includes every SC Medicaid enrollee, who was diagnosed with ADHD at any age between 3 and 25 years old in 2003-2012. Data on injuries and medical treatment as presented in per person/year terms.

4.2 Choice set

There are various ADHD drugs available in the U.S. pharmaceutical market. I improve upon the existing literature on the effects of ADHD treatment on child outcomes by constructing a comprehensive choice set relevant for an ADHD patient. I distinguish between meaningfully different pharmacological treatments and include behavioral therapy into the choice set.

To form a choice set, I group all drugs that were approved by the U.S. Food and Drug Administration (FDA) for treatment of ADHD in children into nine choices by active ingredient and release speed. They are listed in the Table 4 with their respective in-sample market shares calculated for the entire period between 2003 and 2012. The last category, "Others" includes medicines that had an in-sample market share lower than 5%. The market is dominated by the extended-release formulations of relatively old drugs: together amphetamine salts and methylphenidate comprise almost a half of the market for ADHD pharmacological treatments. Stimulants are often recommended as the first step in treatment. Note that most stimulant drugs had seen their patent expire, and there are generic substitutes available in the market.

Although not approved for the treatment of ADHD, certain antidepressants and sleep-disorder medications are prescribed to patients off-label. For example, Provigil (sleep disorders); Wellbutrin (antidepressant); tricyclic antidepressants; Catapres and Tenex (short-acting forms of high blood pressure medicines); Abilify, Zyprexa, Seroquel, Risperdal, and Geodon (antipsychotics). My data allow for identifying off-label prescription practices. Accounting for the off-label treatment is one of the possible extentions of my work.

The provision of prescription drug coverage is an optional benefit that is currently offered by all states. In SC, children and young adults (under age 19) face zero copayment for the prescription drugs and are only responsible for the pharmacy dispense charge (about $5). The state maintains preferred drug lists for medicines that do not require prior authorization. All other drugs may be covered if a doctor-filed authorization request is approved. The quantity restrictions are also common with a typical prescription capped at a 30-day supply.

Besides pharmacological treatment, patients can benefit from behavioral therapy. Therapy usually consists of educating parents and teachers on how to provide positive feedback on desired behaviors and how to discourage unwanted behaviors. Behavioral therapy alone was found to be less effective than pharmacological treatment alone, but no consensus exists on whether medications are inferior to the combination treatment (Barkley, 2006). A combination of behavioral therapy and pharmacological treatment constitutes yet another choise in the set.

A case when a child is diagnosed with ADHD but receives no treatment is the final option in the choice set. Even if treatment is not administered, having their child diagnosed with ADHD is likely to affect her parents' behavior. They might invest their time into development of their child's noncognitive skills.[11]

Table 4: Choice Set in the U.S. ADHD Drugs Market, 2003-2012.

Active Ingredient	Speed	Mkt share	Major Brands	G	Entry	Avg. Price
Amphetamine salts	E	25.16	Adderall XR	Y	2001	150.67
Methylphenidate	E	20.26	Concerta	Y	2000	131.00
Methylphenidate	N	11.13	Ritalin LA, Metadate CD, Methylin ER	Y	2002	127.35
Lisdexamfetamine	E	11.04	Vyvanse	N	2007	141.11
Amphetamine salts	M	8.15	Adderall	Y	1996	37.27
Dexmethylphenidate	E	7.19	Focalin XR	N	2005	144.85
Atomoxetine	n/a	6.37	Strattera	N	2002	130.27
Methylphenidate	M	5.82	Methylin, Ritalin	Y	2002	30.16
Others	–	4.89	Various	–		81.15

Notes: "Speed" stands for the drug release speed, where "E" means extended release, "N" - intermediate and "M" – immediate release speed. Extended release drugs are superior than immediate release drugs in that their active ingredient is released over a longer period of time, often allowing for once-a-day dosing. In-sample market share is based on the number of prescriptions filled in 2003-2012. "G" stands for generic drugs availability. Average price is calculated by averaging SC Medicaid reimbursement payments to pharmacies.

5 Results

The results from estimating single-regression equations include the event of first diagnosis, decision on medical treatment, and four estimations of the effect of treatment on adverse outcomes.

[11] There are no over-the-counter treatments and all major medications from the choice set are covered under SC Medicaid. It means that the claims data contains a complete treatment history.

Table 5: First Diagnosis

Selected variables	Single equation, without unobserved heterogeneity							
History of injuries								
Injuries, t-1	0.013	(0.004)	0.018	(0.005)	0.012	(0.004)	0.012	(0.004)
Injuries, t-2			0.019	(0.005)				
Injuries, t-3			0.003	(0.005)				
Individual characteristics								
Male	0.043	(0.003)	0.036	(0.003)	0.044	(0.003)	0.04	(0.003)
Race: Black	-0.041	(0.003)	-0.030	(0.003)	-0.039	(0.003)	-0.045	(0.003)
Race: Other	-0.025	(0.007)	-0.009	(0.009)	-0.021	(0.007)	0.004	(0.011)
Diag: at 3	-0.435	(0.006)	-0.442	(0.008)	-0.437	(0.006)	-0.41	(0.007)
Diag: at 4	-0.237	(0.005)	-0.236	(0.006)	-0.239	(0.005)	-0.22	(0.005)
Diag: at 5	-0.105	(0.005)	-0.099	(0.006)	-0.106	(0.005)	-0.095	(0.005)
Diag: at 7	0.046	(0.005)	0.047	(0.006)	0.047	(0.005)	0.054	(0.005)
Diag: at 8	0.059	(0.005)	0.054	(0.007)	0.059	(0.005)	0.064	(0.006)
Diag: at 9	0.063	(0.006)	0.056	(0.007)	0.065	(0.006)	0.069	(0.007)
Diag: at 10	0.055	(0.006)	0.038	(0.008)	0.055	(0.006)	0.059	(0.007)
Diag: at 11	0.030	(0.007)	0.008	(0.009)	0.029	(0.007)	0.037	(0.008)
Diag: at 12	0.024	(0.007)	-0.003	(0.010)	0.023	(0.007)	0.019	(0.01)
Diag: at 13	0.028	(0.008)	-0.010	(0.011)	0.026	(0.008)	0.023	(0.012)
Diag: at 14	0.037	(0.008)	-0.016	(0.011)	0.032	(0.008)	0.05	(0.014)
Diag: at 15	0.088	(0.009)	0.037	(0.012)	0.081	(0.009)	0.099	(0.018)
Diag: at 16	0.130	(0.010)	0.093	(0.013)	0.121	(0.010)	0.169	(0.023)
Diag: at 17	0.189	(0.011)	0.132	(0.016)	0.178	(0.011)	0.267	(0.033)
Diag: at 18	0.242	(0.015)	0.211	(0.020)	0.222	(0.015)	0.407	(0.063)
Diag: at 19	0.265	(0.021)	0.254	(0.028)	0.24	(0.021)		
Diag: at 20	0.243	(0.028)	0.254	(0.043)	0.217	(0.028)		
Family characteristics								
N adults					-0.017	(0.001)	-0.017	(0.001)
N children					-0.002	(0.002)	-0.014	(0.002)
Mother's characteristics								
Age							0.0004	(0.0003)
Educ: Less than HS							-0.014	(0.010)
Educ: Some HS							-0.006	(0.008)
Educ: HS diploma							-0.012	(0.007)
Educ: Some college							0.014	(0.008)
Educ: Grad school							0.022	(0.080)
Selected physician types								
Mental/rehab					0.110	(0.026)	0.099	(0.031)
Medical clinic					0.086	(0.026)	0.100	(0.027)
Physician					0.105	(0.026)	0.074	(0.027)
County F.E.	No		No		Yes		Yes	
N obs.	**140,063**		**83,649**		**140,036**		**91,661**	

Notes: The coefficients are marginal effects at means and standard errors are in parentheses. "White" is an omitted chategory in individual race, age of 6 is omitted from individual age, and college-level education category is an omitted category from mother's education. Fourty six county coefficients are not shown.

Table 6: Adverse Outcome: Teen Pregnancy

Selected variables	**Single equation, without unobserved heterogeneity**					
Treatment history						
Treatment, t-1	-0.005	(0.002)	-0.004	(0.001)	-0.002	(0.001)
Treatment, t-2	0.002	(0.002)				
Treatment, t-3	-0.001	(0.001)				
Individual characteristics						
Race: Black	-0.0002	(0.001)	-0.001	(0.001)	-0.0004	(0.001)
Race: Other	-0.009	(0.002)	-0.009	(0.002)		
Pregnant at 11						
Pregnant at 12	-0.037	(0.004)	-0.032	(0.003)	-0.009	(0.002)
Pregnant at 13	-0.037	(0.004)	-0.038	(0.004)	-0.011	(0.002)
Pregnant at 15	-0.014	(0.003)	-0.011	(0.002)	0.002	(0.002)
Pregnant at 16	-0.007	(0.002)	-0.002	(0.002)	0.005	(0.002)
Pregnant at 17	0.005	(0.002)	0.007	(0.002)	0.01	(0.002)
Pregnant at 18	0.009	(0.002)	0.012	(0.002)	0.012	(0.002)
Pregnant at 19	0.017	(0.002)	0.015	(0.002)	0.012	(0.003)
Family characteristics						
N adults			-0.0001	(0.001)	0.001	(0.001)
N children			-0.003	(0.001)	-0.001	(0.001)
Mother's characteristics						
Age					-0.0002	(0.0001)
Educ: Less than HS					0.002	(0.003)
Educ: Some HS					-0.001	(0.002)
Educ: HS diploma					-0.001	(0.002)
Educ: Some college					-0.002	(0.003)
Educ: Grad school						
County F.E.	No		Yes		Yes	
N obs	**19,451**		**29,321**		**13,827**	

Notes: The coefficients are marginal effects at means and standard errors are in parentheses. "White" is an omitted chategory in individual race, age of 14 is omitted from individual age, and college-level education category is an omitted category from mother's education. Fourty six county coefficients are not shown.

Table 7: Adverse Outcome: Injuries

Selected variables	**Single equation, without unobserved heterogeneity**					
Treatment history						
Treatment, t-1	0.096	(0.004)	0.094	(0.002)	0.105	(0.003)
Treatment, t-2	0.022	(0.005)				
Treatment, t-3	0.004	(0.004)				
Individual characteristics						
Male	0.026	(0.003)	0.031	(0.002)	0.032	(0.003)
Race: Black	-0.057	(0.003)	-0.049	(0.002)	-0.056	(0.003)
Race: Other	-0.075	(0.009)	-0.061	(0.006)	-0.051	(0.009)
Age 3	0.010	(0.013)	0.002	(0.004)	0.000	(0.005)
Age 4	0.055	(0.021)	0.044	(0.014)	0.045	(0.016)
Age 5	0.022	(0.021)	-0.007	(0.008)	-0.003	(0.009)
Age 7	-0.004	(0.015)	0.004	(0.005)	0.002	(0.006)
Age 8	-0.005	(0.014)	-0.002	(0.004)	0.000	(0.005)
Age 9	-0.006	(0.013)	-0.001	(0.004)	0.002	(0.005)
Age 10	-0.016	(0.013)	0.000	(0.004)	0.000	(0.005)
Age 11	-0.009	(0.013)	0.005	(0.004)	0.007	(0.005)
Age 12	-0.010	(0.013)	0.006	(0.004)	0.008	(0.006)
Age 13	-0.010	(0.013)	0.007	(0.005)	0.014	(0.006)
Age 14	-0.005	(0.013)	0.012	(0.005)	0.023	(0.006)
Age 15	0.001	(0.013)	0.015	(0.005)	0.019	(0.007)
Age 16	-0.014	(0.014)	0.004	(0.005)	0.014	(0.009)
Age 17	0.012	(0.014)	0.021	(0.006)	0.027	(0.010)
Age 18	0.010	(0.014)	0.005	(0.006)	0.021	(0.013)
Age 19	-0.054	(0.015)	-0.050	(0.007)	-0.057	(0.021)
Age 20	-0.081	(0.018)	-0.079	(0.012)		
Family characteristics						
N adults			-0.004	(0.001)	-0.008	(0.001)
N children			0.000	(0.001)	0.002	(0.002)
Mother's characteristics						
Age					-0.001	(0.0002)
Educ: Less than HS					0.034	(0.008)
Educ: Some HS					0.021	(0.007)
Educ: HS diploma					0.012	(0.006)
Educ: Some college					0.009	(0.007)
Educ: Grad school					0.002	(0.09)
County F.E.	No		Yes		Yes	
N obs	**96,426**		**199,284**		**124,745**	

Notes: The coefficients are marginal effects at means and standard errors are in parentheses. "White" is an omitted chategory in individual race, age of 6 is omitted from individual age, and college-level education category is an omitted category from mother's education. Fourty six county coefficients are not shown.

6 Discussion

References

High school girls with adhd. http://chesapeakeadd.com/adhd-articles/high-school-girls-with-adhd/, Accessed on May 16^{th}, 2014.

L Eugene Arnold. Sex differences in adhd: conference summary. *Journal of abnormal child psychology*, 24(5):555–569, 1996.

Russell A Barkley. *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. Guilford Press, 2006.

A. Chorniy and L. Kitashima. The effect of adhd medications on health and behavioral outcomes. *Working Paper*, 2014.

Gabriella Conti and James J Heckman. *Economics of Child Well-Being*. Springer, 2014.

Gregory S Crawford and Matthew Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007.

Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

Janet Currie, Mark Stabile, and Lauren Jones. Do stimulant medications improve educational and behavioral outcomes for children with adhd? *Journal of Health Economics*, 37: 58 – 69, 2014.

Soren Dalsgaard, Helena Skyt Nielsen, and Marianne Simonsen. Consequences of {ADHD} medication use for children's outcomes. *Journal of Health Economics*, 37:137 – 151, 2014.

Michael Dickstein. Efficient provision of experience goods: Evidence from antidepressant choice. *Manuscript*, 2014.

Lawrence B Finer and Mia R Zolna. Unintended pregnancy in the united states: incidence and disparities, 2006. *Contraception*, 84(5):478–485, 2011.

James J. Heckman, Jora Stixrud, and Sergio Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24 (3):411–482, 2006.

Ronald C Kessler, Patricia A Berglund, Cindy L Foster, William B Saunders, Paul E Stang, and Ellen E Walters. Social consequences of psychiatric disorders, ii: Teenage parenthood. *American Journal of Psychiatry*, 154(10):1405–1411, 1997.

Thomas A Mroz and Timothy H Savage. The long-term effects of youth unemployment. *Journal of Human Resources*, 41(2):259–293, 2006.

Robert J Resnick. Attention deficit hyperactivity disorder in teens and adults: They don't all outgrow it. *Journal of clinical psychology*, 61(5):529–533, 2005.

Tanja Saxell. Private experience and observational learning in pharmaceutical demand. *Manuscript*, 2013.

STD Fact Sheet. Std trends in the united states.

Andrine Swensen, Howard G Birnbaum, Rym Ben Hamadi, Paul Greenberg, Pierre-Yves Cremieux, and Kristina Secnik. Incidence and costs of accidents among attention-deficit/hyperactivity disorder patients. *Journal of Adolescent Health*, 35(4):346–e1, 2004.

Gabrielle Weiss and Lily Trokenberg Hechtman. *Hyperactive children grown up: ADHD in children, adolescents, and adults* . Guilford Press, 1993.

Zhou Yang, Donna B Gilleskie, and Edward C Norton. Health insurance, medical care, and health outcomes a model of elderly health dynamics. *Journal of Human Resources*, 44(1): 47–114, 2009.