Mental Disorders' Impact on Labor Market Outcomes: Theory and Evidence from ADHD

A Thesis Presented to

The Honors Tutorial College

Ohio University

In Partial Fulfillment of the Requirements for Graduation

from the Honors Tutorial College with the degree of

Bachelor of Science in Economics

Joseph Hartge

April 2015

Contents

Abstract

This thesis contributes to the existing empirical research on Attention Deficit Hyperactivity Disorder's (ADHD) influence on labor market outcomes. Two metrics, earnings and job termination, are used to measure labor market outcomes, and the latter has not been tested in the existing literature. In addition to using ADHD as a binary variable, the age of diagnosis is employed to capture a fuller effect of having the disorder in the labor market. Overall, results argue that agents who are diagnosed between ages 5 and 14 can completely overcome any negative effect that ADHD has on earnings, on average, compared to agents who do not have ADHD. Being diagnosed after age 14, however, yields on average lower wages compared to non-ADHD agents. Interestingly, results also argue that an earlier age of diagnosis increases the agent's probability of being fired compared to a later diagnosis.

Acknowledgments: I sincerely thank Dr. Patricia Toledo for her extraordinary patience and flexibility as she advised this thesis. Without her steadfast dedication to teaching and producing quality research, the experience writing this thesis would not have been so fulfilling. I owe her a great debt of gratitude for encouraging me to pursue this research as my thesis. I also thank Dr. Tia McDonald for her helpful comments along the way. Finally, I have to thank the Department of Economics for its unfailing willingness to provide assistance and the necessary accommodations for this research.

1 Introduction

Among the numerous factors that can influence someone's outcome in the labor market, mental disorders have developed their own niche within the economics literature. Questions center around how mental disorders such as depression, panic disorder, and anxiety influence traditional measures of labor outcomes like earnings. Baldwin and Marcus (2007) were the first to present nationally representative estimates of unexplained outcomes differentials using Oaxaca (1973) decomposition techniques. They find results consistent with others' (Baldwin 1999; Yelin and Cisternas 1997), showing that people with mental disorders earn lower wages and have lower employment rates, on average, compared to those without mental disorders or those with physical disorders. They also find that there is a spectrum of mental disorders which are received differently in the labor market. Specifically, when compared to those who have only physical impairments or no impairments at all, adjustment disorders have more favorable outcomes, albeit still negative outcomes, compared to people with psychotic disorders.[1] So while mental disorders send negative signals or produce poor outcomes for people in the labor market, some might be worse than others.

One mental disorder in particular that has garnered special attention is Attention Deficit/Hyperactivity Disorder (ADHD) which the American Psychiatric Association (2013) defines in the Diagnostic and Statistical Manual of Mental Disorders (DSM - V) as a neurodevelopmental disorder characterized by broadly defined behavioral symptoms at various stages of life. Common associations with the inattentive side

[1] Baldwin and Marcus (2007) explains that "adjustment disorders (e.g., panic disorder, post-traumatic stress disorder) are characterized by a significantly more difficult adjustment to a life situation than would normally be expected." She also explains that "psychotic disorders (e.g., schizophrenia, manic-depressive psychosis) are characterized by delusions (false beliefs) and/or hallucinations (false perceptions)."

of the disorder include difficulty sustaining attention in tasks, prone to distraction by extraneous stimuli, and absent-mindedness. The hyperactive-impulsive symptoms include fidgeting, interrupting others or excessive talking, and impatience in group activities.

The cause of ADHD is up for debate in the psychology literature. Some professionals (Timimi 2004) feel the disorder is a social construct used by adults to explain children's behavior that does not fit their expectations. Others (Polanczyk 2007), however, argue that ADHD is a legitimate mental disorder even though biological evidence has not been found to identify it unequivocally.

Regardless, there is growing evidence that broader economic forces are influencing the diagnosis of the disorder. Hence, economists have started taking a stronger interest in ADHD. For instance, there is economic evidence that a state's education policy can influence the diagnosis rates of ADHD (Bokhari, Schneider 2011; Schneider, Eisenberg 2006).[2] While ADHD has been contentious in the public eye, one thing is apparent - the diagnosis rate continues to increase. From 2001 to 2010, the diagnosis rate for children aged 5 to 11 increased 24 percent, and the increase was higher for blacks than whites, for example, at 57 percent versus 19 percent increase, respectively (Getahun et al. 2013). Furthermore, the increase in the diagnosis rate was substantially higher for children in households with income greater than or equal to $70,000 (Getahun et al. 2013). Descriptive statistics like these alone demonstrate that there might be economic incentives influencing the apparent prevalence of the disorder.

Beyond the curious trend in ADHD diagnosis, there is evidence that the economic cost is substantial. The costs even carry over to others associated with an ADHD

[2] Specifically, the authors find that states with school accountability laws, which tie school funding to demonstrable evidence of student progress, have a higher prevalence of ADHD and the drugs used to treat the disorder.

person. Breining (2014) finds evidence that ADHD is a negative externality on siblings since it diminishes their education outcomes. Literature reviews show that not only are there substantial direct costs for patients and their families, but also there are costs related to comorbidities, criminality, and productivity all of which factor into the economic analysis of ADHD (Matza et al. 2005). In general, estimates argue ADHD costs $143 billion to $266 billion dollars in total for the numerous parties it affects including healthcare, productivity, education, and the justice system (Doshi et al. 2012). These estimates, however, might understate the true cost of the disorder due to data limitations and the difficulty in capturing certain latent costs associated with mental disorders.

Common treatments for the disorder center around drugs known as psycho-stimulants which help mitigate the distracting effects of ADHD, inducing sharper, deeper focus from the consumer. By reviewing the recent psychology literature, Langberg and Becker (2012) find evidence that long-term medication use can improve school grades and decreases the chances of repeating a grade or being absent from school. One paper in particular, Scheffler et al (2009), shows that among children who have ADHD, those who are medicated in elementary school tend to perform better on standardized tests than those who are not medicated.

There has been research into how the disorder influences human capital development since one of its most notorious effects is inhibiting children's ability to focus on school work. Currie and Stabile (2006) test how childhood ADHD symptoms might influence academic and behavioral outcomes. At the baseline ordinary least squares (OLS) models of Canadian and American youth samples, they find positive and significant relationships between ADHD symptoms and delinquent behavior, chances of grade repetition, and chances of special education enrollment. They find negative and significant relationships between ADHD symptoms and math and reading scores. Furthermore, they find that an increase in the severity of the disorder can more than compensate for any positive effects from additional household income, demonstrating how insidious ADHD can be for people's development.

Fletcher and Wolfe (2008) extend Currie and Stabile's (2006) findings with a different data set, the National Longitudinal Study of Adolescent Health (Add Health), somewhat different outcomes, and by shifting the analysis to a somewhat later stage in life. Their baseline OLS results indicate there is a negative and significant impact between ADHD symptoms and grade point average (GPA), years of education, and chances of attending college. They find positive and significant relationships between ADHD symptoms and chances of school suspension, expulsion, and dropout. Also, they find evidence that ADHD might exhibit negative externalities on siblings' education outcomes.

A recent paper from Fletcher (2014) is one of the first to delve into how childhood ADHD influences adult labor market outcomes such as earnings, employment status, and social assistance receipt. Using the Add Health dataset, his tests show that on average people who are diagnosed with ADHD as a child earn less on average, are less likely to be employed as an adult, and are more likely to receive social assistance than those who were not diagnosed with ADHD.

This thesis posits the following hypotheses regarding ADHD's impact on labor market outcomes:

1. Hypothesis 1: ADHD causes agents to earn less, on average, than non-ADHD agents.
2. Hypothesis 2: ADHD increases the probability that an agent gets fired, on average, compared to non-ADHD agents.

3. Hypothesis 3: A one year decrease in the age of ADHD diagnosis mitigates the negative wage differential from Hypothesis 1, on average.
4. Hypothesis 4: A one year decrease in the age of ADHD diagnosis reduces the agent's chances of being fired, on average.

Hypothesis 1 has been tested in the literature (Fletcher 2014), but the earnings model presented here offers modifications that have yet to be discussed in the literature. For example, previous research on the ADHD's economic impact has only controlled for the disorder as a binary variable. This thesis introduces the age of diagnosis in addition to the ADHD dummy variable to model the effect of ADHD. The job termination metric in Hypothesis 2 has yet to be tested against ADHD in the literature. By observing the disorder's impact on the probability of getting fired, I hope to offer a more complete understanding of its costs in the labor market since job termination arguably inflicts emotional costs or, more generally, latent costs.

In addition to these basic hypotheses regarding outcomes, Hypotheses 3 and 4 contribute an analysis of the age of diagnosis' impact on labor outcomes. I expect there to be a nonlinear relationship between the age of diagnosis and earnings. That is, I expect an early diagnosis to help compensate the disorder's negative effect on earnings whereas a late diagnosis will not significantly reduce this gap. Hence, not only do those people with ADHD earn less on average compared to non-ADHD people, but also those diagnosed later in life earn less than those diagnosed earlier. Similarly, I expect ADHD to have a nonlinear effect on the probability of job termination. In general, I anticipate ADHD's negative effect on outcomes to diminish nonlinearly as the age of diagnosis decreases. The diagnosis of any disease or disorder transmits a tremendous amount of information to the patient which influences his decision-making. Studying the age of diagnosis in the context of the economics of ADHD can help develop a better understanding of the value of that information.

The thesis proceeds as follows: In the next section, I present a simple principal-agent model and a model of human capital investment to motivate the empirical work in later sections. Section III presents a description of the dataset and key variables used in the analysis, and it introduces the empirical models used to test the hypotheses. Section IV presents the results of those estimated models. Finally, Section V concludes with a discussion of the results.

2 Theoretical Framework

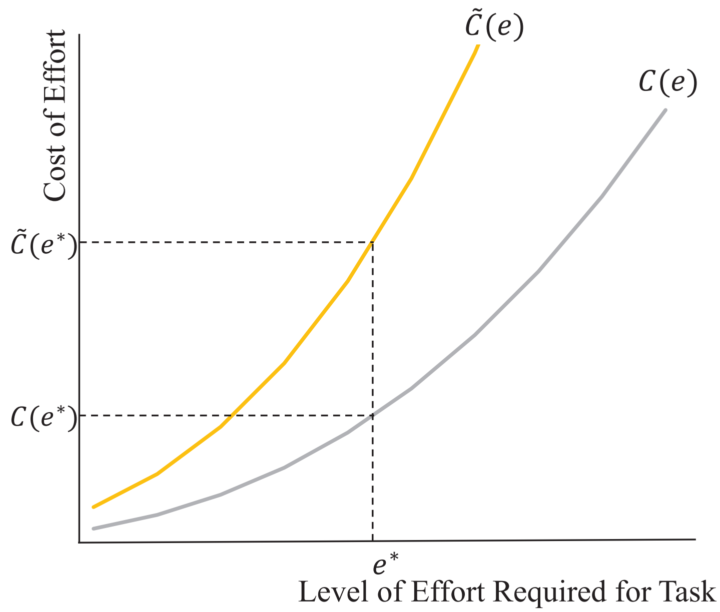
2.1 Principal-Agent Approach

To discuss how ADHD might manifest negative outcomes in the labor market, I use the standard principal-agent model where a principal wants to induce a level of effort, e, for a job which she assumes will cost the agent according to his cost function, $C(e)$.[3]

In this model the worker with ADHD is aware of his disorder. Moreover, the cost function, $\widetilde{C}(e)$, represents the effort that an ADHD agent experiences, and he is aware of his own cost function. In this model, not only is $\widetilde{C}(e)$ not equal to $C(e)$, but also $\widetilde{C}(e) > C(e)$ as illustrated in Figure 1 below.

[3]Conventionally, the agent is male and the principal is female.

Figure 1: ADHD vs. Non-ADHD Agent Cost Functions



Considering the DSM - V (2013) includes symptoms of ADHD such as "failure to pay close attention to details, difficulty organizing tasks and activities, or inability to remain seated in appropriate situations," modelling the ADHD agent's higher cost function in this manner seems to fit well with the disorder itself. Furthermore, in a review of ADHD literature, Wehmeier et al (2010) find that there are specific impairments associated with ADHD that can impact the patient's transition from adolescence into adulthood. Specifically, persistent inattention and emotional impairments associated with ADHD can lead to "poorer work performance in employment settings." They also point out that "impaired planning, anticipation, and preparatory behavior are likely to result in the adolescent not being ready for the future as it arrives" (2010). Thus, someone who has ADHD as a child can carry these impediments as they grow to become an agent in the labor market. Overall, this leads to a generally higher cost of effort compared to an agent who did not have to cope with ADHD as he developed.

I assume that the principal is not aware of the agent's disorder; therefore, she uses the cost function, $C(e)$, when she decides the expected wages to pay the agent. In standard principal-agent models, the principal offers the agent incentives to work that consist of a base salary, α, and a bonus incentive, β, which is a function of the agent's effort and random noise denoted as x. Typically, this unobserved x component captures random idiosyncrasies that create different levels of incentives in the contract market so they each have an expected value of zero and a variance. Overall, this principal's offer is represented as a linear contract for simplicity and takes the form $w = \alpha + \beta(e + x)$. Again, though, the idiosyncratic influence of x has an expected value of zero. This means $E(w) = \alpha + \beta e$.

Assuming rationality, the principal receives some benefit or expected payoff, $P(e)$, from entering the contract with the agent. Thus, her total expected profit from the contract is

$$\pi = P(e) - (\alpha + \beta e). \tag{1}$$

Now, if we assume the agent is risk averse, meaning his marginal utility of wealth diminishes as the terms of the contract become riskier, we can assign him a utility function of the form $U(e) = -e^{-r[w-C(e)]}$ where r is his constant absolute risk aversion and $C(e)$ is his cost to exert the effort to fulfill the contract. Note that this model assumes the principal uses the non-ADHD cost of effort, $C(e)$, when determining her expectations since she is unaware of the agent's ADHD.

Given the fact that the principal assumes $C(e)$ rather than the ADHD agent's true cost, $\widetilde{C}(e)$, it can be shown that the agent's expected utility from effort is

$$E(U(e)) = \alpha + \beta e - C(e) - 1/2 r\beta^2 Var(x). \tag{2}$$

For the agent to accept the contract, the principal must offer compensation to satisfy his reservation utility and so that the level of effort induced by the principal equals the level that maximizes the agent's utility. We will call the first condition, satisfying the agent's reservation utility, the agent's participation constraint (PC) which is defined as,

$$E(U(e)) = \alpha + \beta e - C(e) - 1/2 r\beta^2 Var(x) \leq \bar{U}. \tag{3}$$

So it is clear that this expected utility must be at least as great as the agent's reservation utility lest the agent refuses to accept the contract. Since we can assume the principal's objective is to maximize expected profit,

$$\max_{e,\beta}\{P(e) - (\alpha + \beta e)\}, \tag{4}$$

we will be able to assume that the PC holds as an equality which will be important for the principal to derive the optimal intensity of the incentive, β, to offer.

We will call the second condition, inducing effort that optimizes the agent's utility, the incentive constraint (IC). The principal wrongly assumes that the ADHD agent maximizes effort as the following:

$$e \in argmax_{e^*}\{E(U(e^*)) = \alpha + \beta e^* - C(e^*) - \frac{1}{2} r\beta^2 Var(x)\}. \tag{5}$$

Thus, the optimal level of effort to maximize utility is $\frac{\partial E(U(e))}{\partial e^*} = \beta - C'(e^*) = 0$. Hence, the marginal benefit provided from the extra incentive β equals the non-ADHD agent's marginal cost to exert the effort.

To derive the optimal incentive, β^* , which will induce the level of effort that the principal wants, one notices that agent will only accept the contract if its expected utility is at least equal to his reservation utility, $\bar{U}$. Since the principal is profit-maximizing, however, she will only make an offer that is exactly equal to the agent's reservation utility for the given level of effort. Thus, the PC holds as an equality. From here, it is clear that

$$\alpha + \beta e = \bar{U} + C(e) + \frac{1}{2} r\beta^2 Var(x). \tag{6}$$

So this can be replaced into the principal's objective function as follows:

$$f = \max_{e,\beta}\{P(e) - (\bar{U} + C(e) + \frac{1}{2} r\beta^2 Var(x))\}. \tag{7}$$

And since it has been shown from the IC that $\beta = C^{'}(e)$, this can be replaced into the above equation so that the principal maximizes her incentive entirely as a function of effort. Hence, the objective function becomes,

$$f = \max_{e}\{P(e) - (\bar{U} + C(e) + \frac{1}{2} rC'(e)^2 Var(x))\}. \tag{8}$$

Solving for this objective function yields

$$\frac{\partial f}{\partial e} = P^{'}(e^*) - C^{'}(e^*) - rC^{'}(e^*)C^{''}(e^*)Var(x) = 0. \tag{9}$$

Reversing the previous substitution of β for $C'(e)$ in order to solve for β^*, the objective becomes

$$\frac{\partial f}{\partial e} = P^{'}(e^{*}) - \beta^{*} - r\beta^{*}C^{''}(e^{*})Var(x) = 0. \quad (10)$$

Solving for β^* yields

$$\beta^{*} = \frac{(P^{'}(e^{*}))}{(1 + rC^{''}(e^{*})Var(x)).} \quad (11)$$

In the principal's eyes, given the assumed cost of effort, this is the optimal bonus incentive for the principal to offer the agent in order to induce e^*.

Since the principal is still unaware that the agent has ADHD, the optimal bonus derived in equation 11 will be greater than the optimal bonus offered if she knew he had ADHD. Recall, that she offers β^* because she expects to receive e^* in return. By offering β^*, however, this will not satisfy the ADHD agent's true first order condition (FOC) for optimizing the expected utility of effort. Note that the first order condition assumed by the principal is

$$\frac{\partial E(U(e))}{\partial e^{*}} = \beta^{*} - C^{'}(e^{*}) = 0. \quad (12)$$

But when β^* and e^* are replaced in the true objective function,

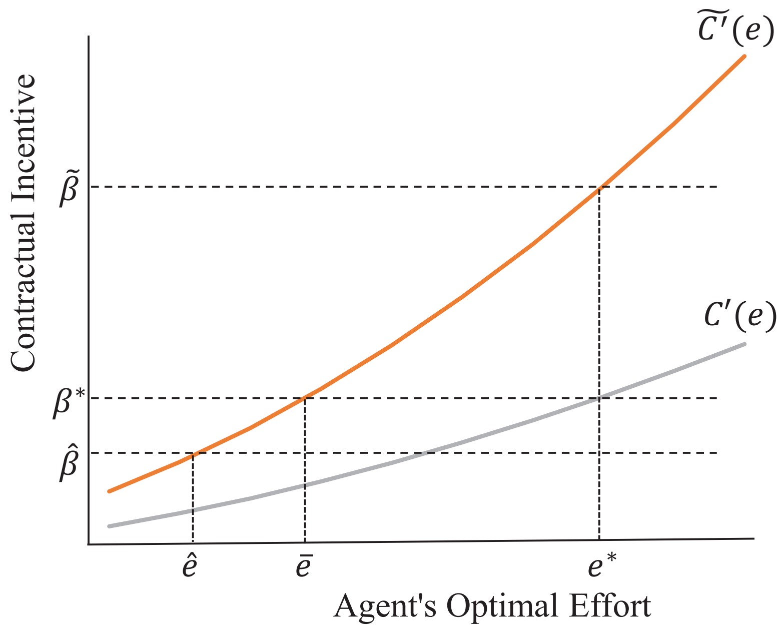
$$\frac{\partial E(U(e))}{\partial e^{*}} = \beta^{*} - \widetilde{C^{'}}(e^{*}), \quad (13)$$

we no longer find that that level of β exactly equals the marginal cost of effort for this ADHD agent since

$$\beta^{*} - \widetilde{C^{'}}(e^{*}) < 0. \quad (14)$$

This disparity between the ADHD and non-ADHD agent's response to production incentives is further illustrated by Figure 2 below, where $\widetilde{\beta} = \beta^{*} + [\widetilde{C^{'}}(e^{*}) - C^{'}(e^{*})]$.

Figure 2: Disparate Effort from Misperceived Costs



This graph helps illustrate how exerting a level of effort, e^*, like the principal wants the agent to do will actually cost the principal a great deal more than she expects to pay since the agent has ADHD. From the example presented in the graph, by offering β^* the principal will incentivize $\bar{e}$ such that

$$\beta^* - \widetilde{C'}(\bar{e}) = 0 \tag{15}$$

where $\bar{e} < e^*$. At this point, we must still assume that the PC holds such that the agent accepts this incentive. Notice, however, that this situation is at the expense of the principal since the marginal benefit of the effort received is less than the marginal cost for the effort, or $P'(\bar{e}) < \beta^*$. Yet she offers β^* because she is unaware that the agent has ADHD and operates under a higher cost function.

While the dispersion of the x noise term can influence this disequilibrium, this model simplifies this by assuming the $E(x) = 0$. Since the principal's marginal cost exceeds her marginal benefit with the ADHD agent, she will have to respond to correct the disequilibrium since she is profit-maximizing. One option might be to offer a new, more optimal bonus incentive to fit the ADHD agent's cost function such as $\hat{\beta}$ in Figure 2, where $\hat{e}$ is profit-maximizing under the new contract. Another option is to fire the ADHD agent and establish more efficient screening mechanisms to prevent her hiring an ADHD agent in the future. Regardless of how the principal chooses to respond, this presents an interesting economic dilemma for dealing with ADHD in the labor market.[4]

2.2 Discussion of the Assumptions

2.2.1 *Rationality of the ADHD Agent*

If an agent has the chance to accept a high-paying contract for which he is under-qualified, it could be perfectly rational for him to accept it to maximize short-run gains despite full knowledge that he is likely to be fired. Not only can he maximize short-run gains, but also he can maximize long-run gains if he selects into a relatively low-paying job after being fired. Overall, this would maximize his total gains. As long as the proper cost assumptions are in place, it is rational for the ADHD agent to accept the non-ADHD contract. For instance, a pragmatic agent like this would have no emotional costs associated with being fired. After all, this is part of his plan to maximize his total gains. The costs related to searching for a job after being fired also would have to be extremely low. Finally, there would have to be no costs incurred to his reputation or signal in the labor market. In reality, however, these costs might not be negligible

[4]Since the principal is profit-maximizing, she will not raise the contractual incentive to $\widetilde{\beta}$, where the ADHD agent would deliver the desired e^*.

which might lead a rational agent not to select a job from which he will be fired.

A myopic agent, on the contrary, might not decide so rationally, and there is evidence that ADHD might inherently make agents more myopic than non-ADHD agents regardless of the potential job termination costs mentioned previously. Wehmeier et al. (2010) present a literature review that explains how "ADHD may involve significant disruption to the brain's executive functioning system which is believed to underlie the human capacity for self-organization and goal-directed actions, or self-regulation [of emotions]." In general, they explain that ADHD often causes emotional impairments in adolescents including "poor self-regulation of emotion, greater excessive emotional expression, [...], problems coping with frustration" and others. Finally, their review shows that for an adolescent with ADHD, "Future rewards are less valued, and so the adolescent shows poor delay of gratification and does not persist toward future goals. Poor inhibition results in poor regulation of emotions, with deficient control of anger and frustration being the most impairing problems in this respect."[5] If the ADHD agent values future incentives less than the non-ADHD agent, he will offer less effort than the principal expects in each contract he faces.[6]

On one hand, this could still be a way for the agent to maximize his overall gains. The agent effectively perfectly price discriminates in the labor market, starting at the highest paying contract and taking each successive lower-paying contract as he descends toward his optimal contract where the β incentive matches the expected effort which equals the agent's optimal effort. For a myopic agent, the costs of being fired (e.g. search costs and diminished reputation) do not matter; therefore, in contrast to a

[5] The idea that ADHD agents might not respond to future rewards as strongly as non-ADHD agents has interesting economic implications about the agents' time preference. This is discussed further in Section 5.

[6] Save the single case where the contract incentivizes the ADHD agent's optimal effort exactly.

rational agent, the myopic agent might systematically make incorrect decisions when selecting into a job.

If the ADHD agent is myopic and continues to select into jobs for which he is not qualified, he will incur costs as he moves from lower-paying contract to lower-paying contract. Costs might include the emotional costs (e.g. increased anxiety, depression, or lack of confidence) of being fired so frequently, job search costs, and diminished value in the labor market.[7] Rather than voluntarily selecting into lower-paying jobs at no cost to him, labor market forces involuntarily move him down toward his optimal contract, incurring costs along the way.

Despite this potentially costly fate of the ADHD agent in the labor market, he might be able to overcome it. This thesis assumes that early diagnosis is highly correlated with early treatment; therefore, if an agent has an early age of diagnosis, the early treatment can cause his cost function to converge to the non-ADHD agent's cost function by the time he enters the labor market, effectively eliminating any earnings gap. Furthermore, the information provided by a diagnosis early in life can help the agent learn how to select utility-maximizing work. That is, an early age of diagnosis can help the agent select the occupation for which he is best suited to optimize his cost of effort and, therefore, earnings.[8] By selecting this optimal occupation, the agent avoids the risk of failure and any frustration that would be associated with selecting a job not suitable to ADHD agents.

[7] In this case, an agent's value in the labor market might be reduced from poor recommendations or a reduced signal or reputation due to gaps in employment and being fired several times.

[8] There are certain occupations such as high-risk jobs that might require less formal education or training but still pay comparable wages to jobs requiring a Bachelor's degree, for instance. The nature of the job offers a wage premium to compensate the agent for the extra risk he bears, thereby eliminating an earnings gap between ADHD and non-ADHD agents. Even still, the nature of this line of work might affect the agent's probability for job termination as is discussed in Section 5.

2.2.2 *Asymmetric Information*

This framework also assumes to an extent that the principal is unaware of the agent's disorder until after they enter the contract. Considering certain policies, such as the American with Disabilities Act of 1990 (ADA), are in place to combat employer discrimination against people with disabilities, this might be a fair assumption. Although, there is significant literature that suggests policies such as these do not improve the employment prospects of disabled people; rather, some would argue it makes it more expensive to hire these people (Acemoglu and Angrist 1998; DeLeire 2000; DeLeire 2001). Thus, legislation that attempts to eliminate wage differentials or employment differentials might exacerbate them to an extent. So even if the agent does not have to reveal his disorder to the principal under these laws, principals and firms in general have screening mechanisms to offset these increased costs associated with new regulations to render disabled agents' employment prospects unchanged.

2.3 Two-Stage Human Capital Investment

The principal-agent model in Section 2.2 illustrates a framework for the decision-making process when the agent enters the labor market. There are important decisions that bring him to that point, though, such as his education and health care choices. As discussed in the previous section, I assume that early diagnosis is highly correlated with treatment. Here, I use a two-stage model of human capital investment to frame the parents' decision to treat their ADHD child.

Since ADHD can be a serious burden for some children's development, families often explore treatment options. One of the most common is medication treatment, although psychological and environment-based treatment options are also used (Goldman et al

1998). The medication involved in treating ADHD works to offset the distracting effects of the disorder, inducing sharper mental focus in the patient. Since ADHD occurs on a spectrum of severity (Hinshaw and Scheffler 2014), however, not all children diagnosed with ADHD might need the medication to function at a satisfactory level.[9] And even if most children with ADHD purchase medication for it, the costs might vary from case to case depending on drug costs, insurance coverage, or other factors. Clearly, families need to make a calculated decision when faced with ADHD.[10]

Using a standard two-stage model of human capital investment, I model the parents' decision to seek treatment for their child. To begin, there are only two periods that make up this decision process. In the first period, the parents earn an income y_i, consume c, save s, and decide whether or not to medicate their child, $m = 1$ or 0, respectively. At the end of period 1, the parents become irrelevant to the outcome. Since we have established that there is heterogeneity among children diagnosed with ADHD (Hinshaw and Scheffler 2014; Scheffler et al 2009), the cost of medication, θ_i, varies with each child i. Finally, in period 2 when the children now have grown into the labor market, those who were medicated receive a wage w_m and those who were not receive w_u. Overall, household utility in each case is given as:

$$U_i = ln(c_i) + ln(\hat{c}_i) \tag{16}$$

where $\hat{c}_i$ is the child's consumption decision, and each household will seek to maximize

[9] Scheffler et al (2009) report that about 56% of children diagnosed with ADHD take medication to treat it.

[10] The dataset used to test the theory in this thesis is introduced in section 3. Unfortunately, it does not provide a convenient variable that accurately accounts for an ADHD respondent's treatment regiment; therefore, the economic impact of treating ADHD is not modeled empirically in this thesis. It would be remiss, however, not to discuss the economics of ADHD medication when analyzing the disorder's labor market impacts. By thinking about this decision in the theoretical framework here, I hope to motivate future research on the treatment of ADHD.

this utility.[11]

In order to maximize its utility, though, the household will have to consider the budget constraints of both the parents in period 1 and the child who will enter the labor market in period 2. I assume that the households have access to credit. Given the concave utility function in equation 16, the parents are subjected to the following budget constraint:

$$c_i + m_i\theta_i = y_i + s \tag{17}$$

where s represents the parents' credit. Equation 17 holds as an equality since the household has a strictly increasing concave utility function, making it preferable for them to consume as much as possible. This leaves the offspring with the following budget constraint in period 2:

$$\hat{c}_i + s(1+r) = m_i w_m + (1-m_i)w_u = w_u + m_i(w_s - w_u) \tag{18}$$

where r is the interest rate on the debt that the offspring is now responsible for since the parents are irrelevant after period 1.

To derive the budget constraint of the household, we solve equation 18 for s and substitute it into equation 17. So

$$s(1+r) = m_i w_m + (1-m_i)w_u - \hat{c}_i$$

$$s = \frac{m_i w_m + (1-m_i)w_u - \hat{c}_i}{(1+r)}. \tag{19}$$

Then, substituting equation 19 into period 1's budget constraint, we get the following

[11] This model is similar to the one found in Drs. Daron Acemoglu and David Autor's (2009) lecture notes at MIT.

budget constraint:

$$c_i + m_i\theta_i = y_i + \frac{m_i w_m + (1 - m_i)w_u - \hat{c}_i}{(1+r)}$$

$$c_i + \frac{\hat{c}_i}{(1+r)} = y_i - m_i\theta_i + \frac{m_i(w_m - w_u)}{(1+r)} + \frac{w_u}{(1+r)}. \tag{20}$$

Again, given the household's concave utility function, these equations hold as equalities. At this point, the household can make a rational decision about medicating the child[12].

To solve for the optimal decision, medicating or not medicating, we maximize the household's utility subject to the budget constraint in equation 20. Since m is not in the objective function (equation 16), the treatment decision can be evaluated using the budget constraint in equation 20 alone.[13] Take the case where $m = 1$ below:

$$c_{i,1} + \frac{\hat{c}_{i,1}}{(1+r)} = y_i - \theta_i + \frac{w_m}{1+r}. \tag{21}$$

Now, consider the case where $m = 0$ below:

$$c_{i,0} + \frac{\hat{c}_{i,0}}{(1+r)} = y_i + \frac{w_u}{1+r}. \tag{22}$$

From this, it is clear that the parents are indifferent to treatment if $c_{i,1} + \frac{\hat{c}_{i,1}}{(1+r)} = c_{i,0} + \frac{\hat{c}_{i,0}}{(1+r)}$ between equations 21 and 22. Setting the right-hand side of each of these

[12]Note that there is only one period of discounting in this simple two-stage model. In an n-period model, we would have to discount the n-1 periods' consumption and earnings (those after period 1) up to $(1+r)^n$.

[13]This is driven by the Separation Theorem which states that human capital accumulation and supply decisions can be *separated* from consumption decisions.

equations equal and solving for θ yields

$$\theta_i = \frac{w_m - w_u}{(1+r)}. \tag{23}$$

From this, it is clear that the parents will only choose to treat their child if the discounted wage differential from the treatment is *strictly greater than* the cost of treatment (i.e. $\theta < \frac{w_m - w_u}{(1+r)}$).[14] In theory, of course, the wage component w can be generalized to represent the full benefits associated with treatment.

2.4 Discussion of Assumptions

2.4.1 *Time-consistency*

As we understand from Becker and Mulligan's (1997) major contribution, someone's consumption decision in the present does not necessarily match their consumption decision in the future. The two-stage model of ADHD treatment presented here, however, does not consider future treatment decisions. In this model, the treatment decision is made in period 1 only.

The information regarding ADHD that is used to make a treatment decision in the present might not be the same when making that decision in period 2, however. Consider the side effects of ADHD treatment, particularly medication treatment. Cascade et al. (2010) explain that there are "relatively common adverse events that may impact, and even impair, short- and long-term outcomes" despite the medications' noted clinical benefits. Common side effects can include but are not limited to loss of appetite and weight, sleep deprivation, and mood swings, and they vary by the age of the patient

[14]If θ is equal to the discounted wage differential, then the parents are indifferent to the treatment.

(Cascade, Kalali, Wigal 2010). The parents might decide to medicate their child in period 1, but after period 1, when the parents become irrelevant, the side effects might be impairing enough for the child (who is now an adult) to decide not to use the medication in period 2. Overall, the optimal treatment decision could change from period 1 to 2.

Whether or not this treatment decision changes from period 1 to 2 could affect the child's labor market outcome. For instance, if the household is time-consistent, then deciding to take medication in both periods 1 and 2 could help the child emulate a non-ADHD agent in the labor market which could maximize earnings and minimize the chances of being fired. Also, side effects incurred in period 1 could be corrected by changing the treatment decision (switching from medication to environment-based treatment, for example) in period 2. By switching the treatment decision in this case, the agent can still minimize the negative impact of having the disorder in the labor market. In general, it is desirable to have a more complete theoretical that incorporates facts crucial to the treatment decision.

2.4.2 *Asymmetric Information*

Another pitfall of this model is its assumption about information. When a parent is calculating the decision to medicate his child or not, his primary source of information, for both the present and the future, is a doctor. If there are any conflicts of interest in that patient-doctor relationship, however, there is information asymmetry that can lead to poor decision making.

Take the case where the doctor has a contract with a brand-name ADHD medication provider and receives payments when he chooses to prescribe that drug to his ADHD patients. Also, assume the proper checks are in place to prevent rampant over-diagnosis by a single doctor. This provides a serious incentive for him to prescribe medication to those who show just enough symptoms to be diagnoses even if they could get by without them. It also encourages him to prescribe the more expensive, brand-name drug rather than a generic drug which could incur unnecessary costs on the patient's family.

By assuming there is perfect information in the decision making process inappropriately disregards other economic considerations such as moral hazard. Further research can focus on exploring this idea and others to help develop a more generalizable model.

3 Dataset Description and Methods

3.1 Dataset and Sample Design

This thesis employs the National Longitudinal Study of Adolescent to Adult Health (Add Health) data set from the University of North Carolina's Carolina Population Center. From 1994 through 1995, the survey initiated a series of in-home questionnaires to a cohort primarily composed of children in grades 7-12 at the time. This first wave of the study also addresses the children's parents and school administration in separate surveys, but the main focus of the data collection throughout all of the years is the cohort of children. Wave II of the survey began in 1996 when the adolescents were in grades 8-12. Wave III covered 2001 through 2002 when the now young adults were aged approximately 18-26 years. Finally, Wave IV covered the years 2007 through 2008 when the now adults were aged approximately 24-32 years. At each of these four waves, Add Health receives responses from between 14,700-21,000 individuals.

Wave I begins with a clustered sample of high schools from which the Add Health administrators derive respondents of interest. Regarding the sample design, the survey administrators explain that "systematic sampling methods and implicit stratification ensure that the 80 high schools selected are representative of US schools with respect to region of country, urbanicity, size, type, and ethnicity" (Harris 2009). From this sample, an in-school questionnaire was administered to more than 90,000 students in grades 7-12. There was also an in-home sample derived from this same stratified cluster of schools. The corresponding in-home questionnaire is the primary source of data used in this thesis. To create this in-home sample, students in each school were stratified by grade and sex.

3.1.1 *ADHD Metrics*

Regarding the measurement of ADHD, Add Health offers two major methods to track the diagnosis. Firstly, in Wave III Add Health administered a "Retrospective Attention Deficit Hyperactivity Disorder" portion of the in-home survey which Fletcher and Wolfe (2008) employed to measure a respondent's ADHD status.[15] This methodology was highly similar to Currie and Stabile's (2006) ADHD metric which used a sample from the National Longitudinal Survey of Youth (NLSY). Both methodologies derive a 'hyperactivity score' from a series of questions that directly address common symptoms of ADHD as noted by the American Psychiatric Association (DSM - V). Using this method, there is no formal diagnosis of ADHD; rather, the researchers essentially determine the respondent's level of inattention and/or hyperactivity depending on his score. Since ADHD diagnosis is notoriously imperfect, there are clear advantages to

[15]This portion of the in-home survey tells the respondents (about 18-28 years old at this point) to "think back to when you were between 5 and 12 years of age." Then it asks a series of 18 questions about behavior commonly associated with ADHD and how closely they related to that behavior at that age.

using this method to try to capture the best sample of ADHD subjects as possible. Furthermore, it helps control for any relevant moral hazard issues that might incentivize physicians' misdiagnosing the disorder. One particular advantage to the Add Health data using this methodology is that the responses are self-reported whereas the NLSY variables are parent-reported. The disadvantage is that the Add Health data are retrospective whereas the NLSY data are not. Even though there are some advantages to using a scaled score like the one available from NLSY or Wave III of Add Health, there are many more moving parts inherent with this ADHD metric which leave it vulnerable to measurement error. A simpler alternative would be a variable that explicitly asks about a formal ADHD diagnosis.

The second method that Add Health offers to track ADHD diagnosis offers exactly this alternative. During Wave IV of Add Health's in-home interview (when respondents are about 24-32 years old), respondents are asked, "Has a doctor, nurse or other health care provider ever told you that you have or had: attention problems or ADD or ADHD?" From this binary variable, there is a sample of 775 who respond affirmatively (where yes=1 and means they have received a formal diagnosis). Fletcher (2014) employs this methodology instead of the hyperactivity score available from Wave III. It is important to note that the sample of 775 people who report having ADHD is about 5% of the Wave IV sample which matches extremely closely with general estimates of the prevalence of the disorder (Bloom et al 2010). Despite the imperfections related to ADHD's diagnosis, this is arguably the most direct approach to measuring ADHD for the purposes of this thesis, and the 5% sample size is a positive sign that this sample accurately reflects an ADHD sample.

The very next question in the Wave IV in-home interview asks, "How old were you when the doctor, nurse or other health practitioner diagnosed you with attention problems or ADD or ADHD?" This variable is precisely what we need to test how the age of diagnosis might influence any differential in earnings between the ADHD and non-ADHD subsamples.[16]

To derive the best subsample of ADHD agents as possible, I followed similar methods as Fletcher (2014). For instance, to eliminate the possibility of reverse causation, I eliminated those people who said they were diagnosed in the same year as the Wave IV interview. This prevents people from blaming relatively low annual earnings on an ADHD diagnosis, especially if they have not been officially diagnosed. Furthermore, I eliminated those people who said they were diagnosed in their birth year. Not only is being diagnosed in your birth year impossible, but also this helps control for respondents' self-diagnosis. That is, people who feel they might have attention problems from hearing about the disorder on the news or from friends who have it might be inclined to diagnose themselves for poor productivity rather than seek out a medical professional for confirmation. Since these people never will have received an official diagnosis, they might simply claim they have had the disorder since birth. Even still, this does not control for false age of diagnosis reporting which should be taken into consideration when analyzing the results.

3.1.2 *Earnings Metric*

In wave IV, the respondent is asked, "In [year of interview], how much income did you receive from *personal earnings* before taxes, that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment?" (emphasis added). Thus, the question controls for earnings only, rather than potentially including other income from family members or social assistance programs, for instance. Since the

[16] The 5% response rate for this question aligns directly with the previous question in the survey.

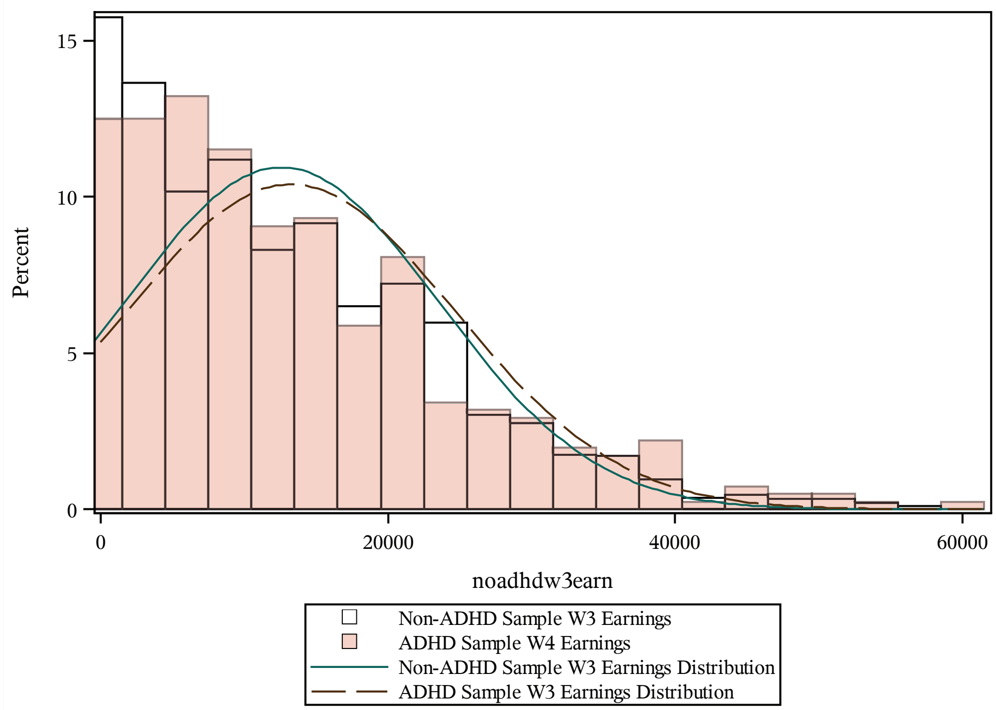
hypothesis regarding ADHD's impact on earnings simply focuses on earnings later in life, this thesis uses wave IV's earnings as a cross-sectional dependent variable rather than constructing a panel of Wave III and Wave IV earnings. Details are given below about why this method of measuring earnings was chosen.

To derive a confident earnings metric from the Add Health data, I only included those who reported some positive earnings. Thus, the 1,069 people who reported $0 in annual earnings during Wave IV were excluded from the sample. As is explained in the next section, both the level of Wave IV earnings and the natural logarithm of this variable is used to measure ADHD's impact.

Figure 3 compares the earnings of the ADHD sample to the entire sample during Wave III.[17] Simply looking at the histograms and respective kernel distributions, one might guess that any difference in earnings between the two samples is negligible. Using the Kilmogorov-Smirnov test to determine whether or not, in fact, the difference is significant, we see that the distributions are statistically equivalent.

[17] The actual earnings question from the Add Health survey refers to the year prior to Wave III interview

Figure 3: Wave III Earnings

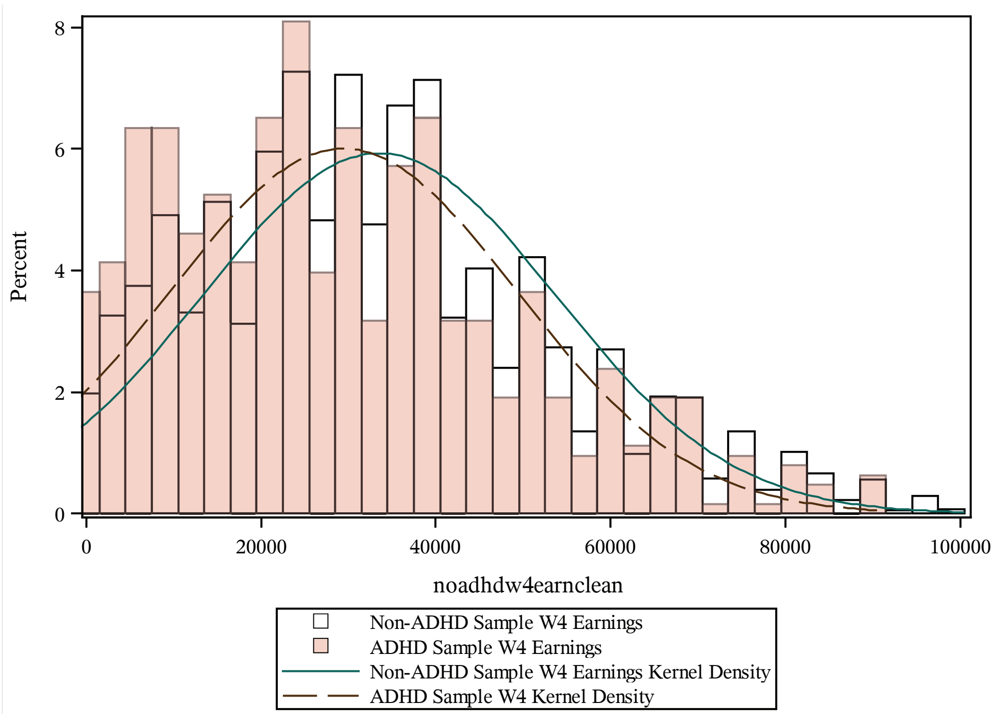


Kolmogorov-Smirnov two-sample test: $p > 0.3126$

The result of the Kilmogorov-Smirnov test is somewhat expected, though, since the average age during Wave III is about 22 years old, meaning the respondents were fairly new in the labor market. The theoretical model proposed in Section 2 suggests that some time ought to pass since first entering the labor market before ADHD agents begin to experience the headwinds of their disorder. Now, we must consider Wave IV's earnings distributions where the average age is about 29 years old.

Looking at Figure 4, a greater earnings disparity seems to appear during Wave IV, fives years after Figure 3. Now, it might be safe to guess that the ADHD subsample earns less than the whole sample on average, and, in fact, the Kilmogorov-Smirnov test confirms this. Considering the results from Fletcher (2014), Figure 4 is expected especially in the context of the theoretical model proposed in Section 2.

Figure 4: Wave IV Earnings

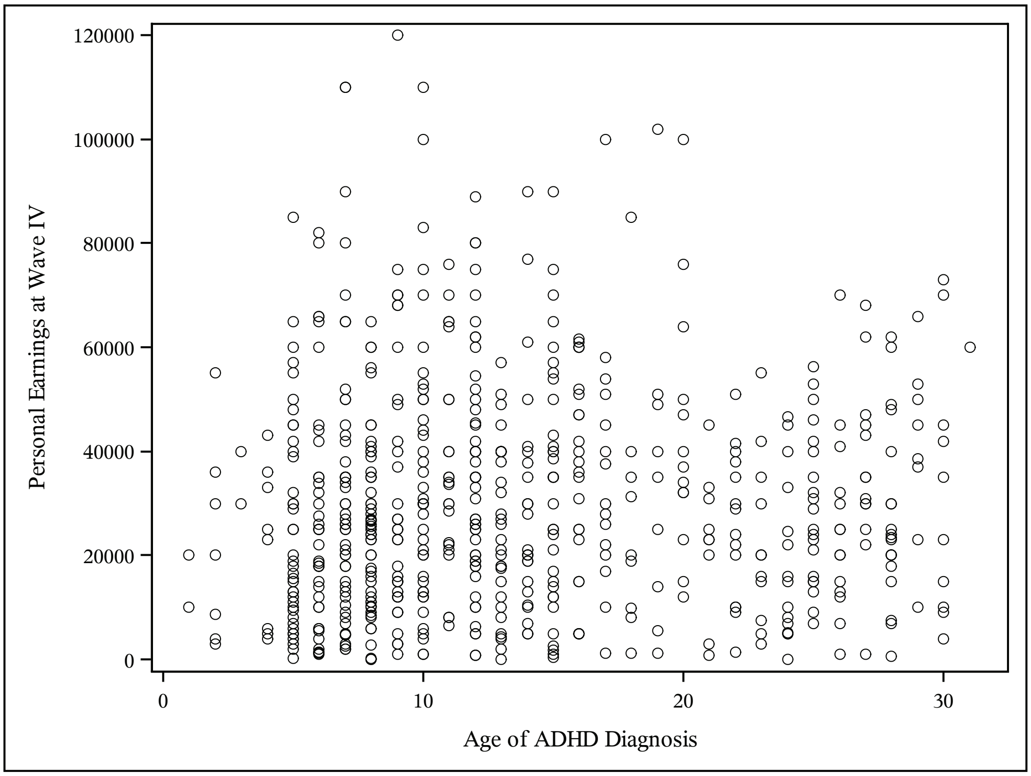


Kolmogorov-Smirnov two-sample test: $p < 0.0001$

Comparing Figures 3 and 4, the theoretical model of ADHD's influence on labor market outcomes would suggest that as time passes in the agent's labor market experience, the principal learns about the agent's disorder (either directly or implicitly from a non-profit-maximizing performance on the agent's end of the bargain) and adjusts her future offerings accordingly. She either renegotiates the contract at a lower β incentive or fires the ADHD agent. After being fired, the agent would select into a more appropriate job that pays less initially compared to the previous job. Thus, ADHD agent's have an especially difficult time advancing in the labor market compared to everyone else.

The relationship between the age of ADHD diagnosis and earnings is another central question of this thesis. Figure 5 below lays out the relationship from the Add Health data.

Figure 5: Age of Diagnosis and Wave IV Earnings



From simple inspection, one can notice a cubic relationship between these variables. Potential earnings at wave IV seems to peak around an age of diagnosis of 12 years old, then trough at about 22 years old, and then peak again at 30 years old. For this reason, the age of diagnosis variable is included in the linear earnings model as a polynomial of degree 3.

3.1.3 *Job Termination Metric*

We need a reliable, precise variable that measures an individual's experience with job termination if we want to include it as a dependent variable in an econometric model. Fortunately, Add Health covers this explicitly in Wave IV. The question asks, "Thinking back over the period from 2001 to the previous year, how many times have you been fired, let go or laid off from a job?"[18] The responses are recorded as a count variable from 0 times up to 50 or more times, and this variable is transformed into a binary variable to be used as the dependent variable in the job termination model.

3.1.4 *Education Metric*

One simple enhancement that this thesis makes to Fletcher's (2014) work is including the respondent's education level in the Mincer model that he estimates. He includes a variable to control for maternal years of education which might be highly correlated with the respondent's education attainment. It would be more preferrable to control for the respondent's education directly, however. The Wave IV in-home interview asks the subject, "What is the highest level of education that you have achieved to date?" The responses include numerous possible education outcomes like "8th grade or less," "some high school," "high school graduate," "completed vocational/technical training (after high school)" and so on as the level of education becomes progressively more rigorous to the level of doctoral and professional degrees.

Using this variable from Wave I, I construct a series of binary variables representing different levels of educational attainment. The variables included are some high school

[18]Note well that nearly all of the Wave IV responses were collected in 2007. Thus, this job termination variable controls for any direct effect of the recession that began right around the time since the question explicitly asks about being fired in the *previous* year, preceding the recession.

experience (but less than a diploma), a high school diploma, some or completed technical training, some college (but less than a degree), bachelor's degree or some graduate school experience, a master's degree or some training beyond a master's or some professional training, and, finally, a completed doctoral or professional degree. The category "eighth grade or less" is also derived from the survey, but it was excluded from the model to avoid multicollinearity.

3.1.5 *Household Income Metric*

In Wave I, a questionnaire was administered to the parents. One question asks, "About how much total income, before taxes did your family receive in 1994? Include your own income, the income of everyone else in your household, and income from welfare benefits, dividends, and all other sources." Using this control for household income in the models helps directly test for the household's investment decision regarding ADHD as explained in section 2.3.

3.1.6 *Delinquency Metric*

The in-home interview questions for Wave I asks the children, "How old were you when you tried marijuana for the first time?" Of the 20,745 responses, 14,606 recorded never having tried marijuana. Most of the balance recorded having tried marijuana from ages 1 to 18.[19] Wave II, which took place less than one year after Wave I, asks the respondents, "Since [month of Wave I interview], have you tried or used marijuana?" There were 3,822 affirmative responses and 10,819 negative responses to this question. This variable was transformed into a binary variable when it was used in the models.

[19] There were a total of 309 responses for either "refused," "don't know," or "not applicable."

Similarly, the Wave I in-home interview asks the children, "Do you ever drink beer, wine, or liquor when you are not with your parents or other adults in your family?" There were 8,405 affirmative responses and 3,190 negative responses. Wave II asks the follow-up question, "Since [month of Wave I interview], did you drink beer, wine, or liquor when you are not with your parents or other adults in your family?" There were 5,379 affirmatives and 1,546 negatives. This variable is included as a binary variable in the models.

Table 1 below displays descriptive statistics for all of the relevant variables described in this section.

Table 1: Descriptive Statistics of Key Variables

	Wave I		Wave IV	
Variable	ADHD	Non-ADHD	ADHD	Non-ADHD
Age	16.14		29.10	
ADHD (n)	520	N/A	123	N/A
Age of Diagnosis (mean)			13.17	N/A
Earnings (mean)	\$4,098.83	\$4,652.92	\$32,666.84	\$37,370.36
Job Terminations (mean)			2.04	1.77
Job Termination (% of subsample)			42.91%	29.76%
Some High School (%)			12.05	7.36
High School Diploma (%)			14.97	16.40
Technical Training (%)			8.74	9.92
Some College (%)			39.47	33.98
Bachelor's Degree (%)			18.01	23.32
Master's Degree (%)			3.44	6.73
Doctoral Degree (%)			1.85	1.91
Adolescent Household Income	\$53,129.39	\$45,728.16		
Black (%)	10.33	23.56		
Hispanic (%)	7.15	16.36		
Female (%)	36.16	54.02		
High School Marijuana (%)	42.91	34.32		
High School Alcohol (%)	56.03	50.06		
Family Social Assistance (%)	24.77	24.13		

Post stratified untrimmed cross-sectional grand sample weight used to compute statistics.

The average age in Waves II and III were 16.81 and 22.37 years, respectively. There are 26 respondents who were diagnosed with ADHD in Wave II and 86 who were diagnosed in Wave III. Average earnings for the ADHD sample in Waves II and III were \$5,584.73 and \$13,538.18, respectively. Average earnings for the non-ADHD sample in Waves II and III were \$5,514.68 and \$13,329.18, respectively.

While no relationships can be directly inferred from the statistics in Table 1, they illustrate some expected results based on the existing literature discussed in section 1. For example, the proportion of ADHD people whose highest level of education is some high school but no diploma is much higher than the non-ADHD sample. The proportion of the ADHD sample whose highest level of education is a high school degree is smaller than that of the non-ADHD sample, however. The same statistics for college does not change this educational attainment story either. The proportion of ADHD people with some college but no Bachelor's degree is higher than that of the non-ADHD sample. The Bachelor's degree statistics, however, show that a higher proportion of non-ADHD people receive a degree compared to ADHD people.

Regarding delinquency, a much higher proportion of the ADHD sample reported using alcohol and marijuana in high school compared to the non-ADHD sample.

The average household income for the ADHD sample was much higher than the average income of the non-ADHD sample which tends to reflect the results explained in Getahun et al. (2013).

Finally, when it comes to labor market outcomes, the average Wave IV earnings for the ADHD sample is lower than the average earnings of the non-ADHD sample. The average number of job terminations for the ADHD sample is higher than that for the non-ADHD sample. Moreover, the proportion of the ADHD sample who experiences job termination is over 13 percentage points higher than the proportion of the non-ADHD sample experiencing job termination.

Many of these relevant variables are captured in the matrices, $\boldsymbol{Z}$ and $\boldsymbol{X}$, in the empirical models introduced in the next section.

3.2 Empirical Models

This thesis employs two types of econometric models to address the three hypotheses. The first is a simple OLS regression which incorporates controls for an ADHD diagnosis and the corresponding age of diagnosis. The second model, however, requires a nonlinear specification so I use a logit model to test this using the dummy variable for job termination as the dependent variable.

Existing literature on the economics of ADHD have only modeled ADHD as a binary variable. By including the age of diagnosis in the models in addition to the ADHD dummy variable, I hope to observe a more revealing effect that the disorder has in the labor market. For such a dynamic diagnosis as ADHD, using only a binary variable to estimate its costs in the labor market is simply too elementary.

3.2.1 *Earnings Model*

I follow Fletcher's (2014) methodology in estimating ADHD's impact on adult earnings by using a traditional Mincer (1974) model, but it also tests the level of earnings, too, which helps illustrate the effect of ADHD more concretely. The Mincer (1974) model was developed by Jacob Mincer in his seminal contribution to the development of human capital theory. It explains earnings, expressed as a natural logarithm, as a function of schooling and a quadratic polynomial of labor market experience. The Add Health dataset does not provide a convenient labor market experience metric so the respondent's age in Wave IV is used as a proxy. The following equation represents the theoretical empirical model of earnings:

$$ln(earnings)_{i,4} = \beta_0 + \sum_{j=1}^{7} r_j S_{i,j} + \beta_1 ADHD_i + \sum_{n=1}^{3} \alpha_n AgeDiag_i^n + \sum_{m=1}^{2} \delta_m Age_{i,4}^m + \boldsymbol{Z\phi} + \epsilon_{i,4} \tag{24}$$

where 4 denotes Wave IV. The r coefficient in the second term represents the rate of return to an additional level of schooling (S), and the Age variables represent the respondent's age at wave IV. The years of schooling variable, S, is a binary variable of the education categories explained in section 3.1.4. Other relevant variables besides ADHD and S are captured in $\boldsymbol{Z}$ and are presented in the Table 1. The additive ϵ term is a classical error term in this linear model.

The ADHD variable is a binary variable where ever being diagnosed with ADHD = 1, and ADHD = 0 otherwise. The age of diagnosis variable ("AgeDiag" in the models) is an interaction term where the ADHD variable is multiplied by the age of diagnosis variable. This way, the age of diagnosis variable takes on a 0 for people without ADHD.

3.2.2 *Job Termination Model*

One of this thesis' primary contributions to the literature is attempting to measure ADHD's impact on the probability of job termination. The following model outlines the relationship:

$$log\Big[\frac{P(Fired=1)_{i,4}}{1-P(Fired=1)_{i,4}}\Big] = \beta_0 + \beta_1 ADHD_i + \beta_2 AgeDiag_i + \boldsymbol{X\phi} + \epsilon_i \qquad (25)$$

Like the earnings model, other relevant variables are captured in $\boldsymbol{X}$.

4 Results

4.1 Earnings Results

Table 2 shows the results from both the Mincer models and level models of earnings.[20] Like Fletcher's (2014) presentation of results, only select regressors are reported here, but a full report of the results is found in Table 4 in the appendix.

The ADHD variable is negative in each model but only significant in Models 1 and 3 where the age of diagnosis is unrestricted. The effect of the age of diagnosis is nonlinear so it is easier to interpret when the predicted values are plotted with the age of diagnosis. We have to consider all four ADHD-related variables simultaneously to understand its effect on earnings.

[20] Although the Mincer models use *ln(HH Income)* as a regressor rather than *HH Income* like the Level models do, both sets of models were tested with both regressors. In the end, the results were robust in both cases.

Table 2: Results of Earnings Models

Regressors	Mincer Models Model 1 Age of Diagnosis > 0	 Model 2 Age of Diagnosis ≥ 5	Level Models Model 3 Age of Diagnosis > 0	 Model 4 Age of Diagnosis ≥ 5
Diagnosed with ADHD	-1.3695^{**} (0.6716)	-1.7511 (1.0902)	-26462^{**} (12138)	-21618 (18735)
Age of ADHD Diagnosis	0.2772^{*} (0.1528)	0.3603 (0.2382)	5645.54^{*} (3008.17)	4599.18 (4312.36)
Age of ADHD Diagnosis^2	-0.0196^{*} (0.0105)	-0.0249 (0.0155)	-407.36^{*} (209.69)	-341.23 (285.61)
Age of ADHD Diagnosis^3	0.0003954^{*} (0.0002118)	0.0004962 (0.0003020)	8.1410^{*} (4.2950)	6.8818 (5.6413)
Age	0.2672 (0.2872)	0.2689 (0.2884)	13620 (9065.45)	13244 (9153.11)
Age^2	-0.0039 (0.0049)	-0.0039 (0.0050)	-201.61 (154.82)	-195.00 (156.36)
ln(HH Income)	0.1470^{***} (0.0296)	0.1474^{***} (0.0296)		
HH Income			54.28^{***} (16.3197)	54.23^{***} (16.3264)
Female	-0.4079^{***} (0.0318)	-0.4088^{***} (0.0319)	-12640^{***} (1056.04)	-12662^{***} (1056.71)
Black	-0.1321^{***} (0.0429)	-0.1346^{***} (0.04296)	-4957.63^{***} (1357.70)	-4925.07^{***} (1362.68)
Hispanic	0.0521 (0.0616)	0.0526 (0.0616)	-9.87 (1432.85)	-12.07 (1433.20)
Adjusted R-squared	0.1363	0.1362	0.0755	0.0753
n	9,850	9,837	9,850	9,837

Heteroskedasticity-robust standard errors reported in parentheses
Post stratified untrimmed cross-sectional grand sample weight used to estimate models.
**** implies estimate is significant at 1%*
*** implies estimate is significant at 5%*
** implies estimate is significant at 10%*

The dependent variable comes from respondent interview in Wave IV when $n = 15,701$, but the household income variable comes from the parent interview in Wave I when $n = 17,670$. This disparity between sample sizes can generate some blanks in the final matrix used to estimate the model. Furthermore, there were a total of 1,847 unusable responses from the earnings variable, 2,447 unusable responses from the household income variable, and 901 unusable entries from the sample weights. Unusable responses might include legitimate skips in the interview questions, refusal to answer, "don't know," or simply missing observations among other possibilities. There were 25 people who were diagnosed before age 5 which leaves Models 2 and 4 with smaller samples than the unrestricted models.

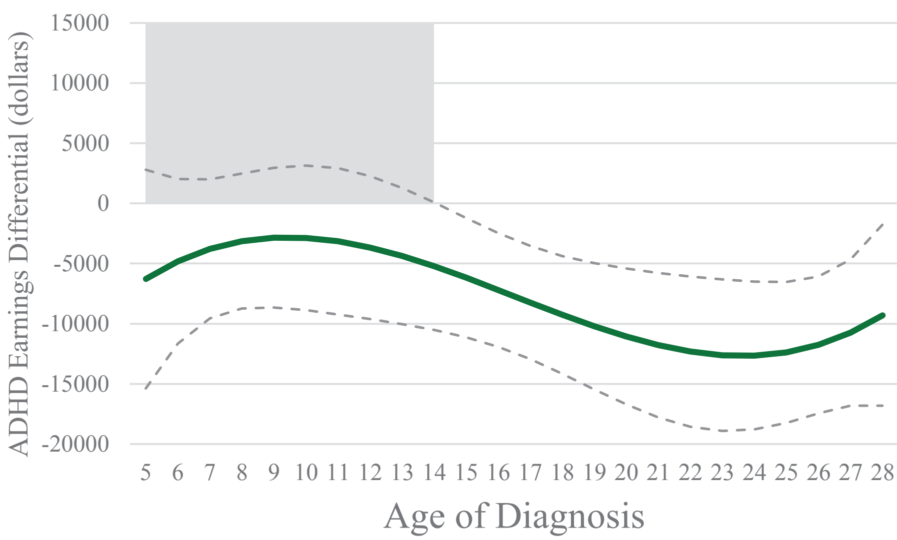
In the preliminary estimation of these models, there were issues with statistical noise on both tails of the age of diagnosis influencing the results. Furthermore, the DSM - IV, which was used by the psychology profession during the Add Health surveys, states that the symptoms for ADHD must have been present by age 6, making all ages leading up to 6 and thereafter eligible for diagnosis. I anticipate that the number of diagnosis increase as the age of diagnosis approaches 6 so I compared the number of diagnoses at each age leading up to age 6. Overall, at ages 1, 2, 3, and 4 there are 3, 8, 4, and 10 responses of an ADHD diagnosis, respectively. At age 5, however, there are 47 responses. To be conservative with my sample yet still accurate by removing outliers, I considered the marginal increase of 37 diagnoses from ages 4 to 5 as an indication that age 5 is a starting point to analyze the age of diagnosis. Thus, Models 2 and 4 place a closed lower bound on the age of diagnosis at 5 years old.

Figure 6 shows ADHD's total effect on earnings as estimated by the Model 4 in Table 2. The slope of the predicted function in Figure 6 is the effect of the age of diagnosis on earnings.[21] Since the slope is generally negative in this graph, the results align with the expectations – increasing the age of diagnosis decreases average wages in Wave IV. From the results in Models 1 and 3, the negative effect on earnings is statistically significant as indicated by the p-value of the estimated coefficient of the cubic variable. The results in Models 2 and 4, however, show that this result is not very robust in terms of statistical significance.

[21] Given equation 24, the slope of Figure 6 is

$$\frac{\partial ln(earnings)}{\partial AgeDiag} = 0.2772 - 2(0.0196)AgeDiag_i + 3(0.0003954)AgeDiag_i^2. \tag{26}$$

Figure 6: Wave IV Earnings Differential for ADHD vs. Non-ADHD Agent



The solid line is ADHD's predicted effect on the average earnings in Wave IV compared to a non-ADHD agent. The dashed lines represent a 95% confidence interval for the prediction. The shaded region indicates that the ADHD agent's earnings are statistically equal to the non-ADHD agent's, on average.

As indicated by the shaded region, there is a range of ages early in life where an ADHD diagnosis still leaves the agent's earnings statistically indistinguishable from a non-ADHD agent's outcomes. For instance, someone who is diagnosed at age 10 earns the same amount in Wave IV as someone without ADHD, on average. The results from Model 4 argue this holds for ADHD agents diagnosed up to and including age 14. A diagnosis at age 15 and beyond, however, yields persistently negative earnings differentials for the ADHD agent. This seems to be early evidence that there is, in fact, an optimal age of diagnosis. For example, when the age of diagnosis equal is 15, the average earnings gap could be as much as $10,000 per year.

Nevertheless, the model's specifications can be improved in further research to derive a more confident, thorough estimate of the effect that the disorder has on labor outcomes.

The results not only for ADHD's effect on earnings, but also all of the regressors' effects on earnings were found to be quite robust as the models were developed. It is especially encouraging that the sign and magnitude of ADHD's effect held up against changing model specifications and samples. Although the statistical significance levels for ADHD in the age-of-diagnosis-restricted models (Models 2 and 4) change from the unrestricted models, the statistical inference from the confidence intervals, as shown in Figure 6, remain largely unchanged. Since the weighted sample of ADHD agents is 755 (which is about 5.06% of the entire weighted sample of agents), restricting the age of diagnosis to a minimum of 5 years old can influence the statistical significance since a larger sample would reduce the standard errors of the estimates.

4.2 Job Termination Results

Table 3 shows the results from the logit model of job termination. Figure 8 in the appendix shows that the job termination and age of diagnosis variables have a similar cubic relationship as earnings and the age of diagnosis. This relationship was tested in the logit job termination model the same way it was in the linear earnings model - with quadratic and cubic age of diagnosis controls. This specification did not yield any viable results, though, so the job termination model was tested only with age of diagnosis of degree one. The relationship illustrated in Figure 8 still helps understand the results of the job termination model.

Table 3: Results of Job Termination Model

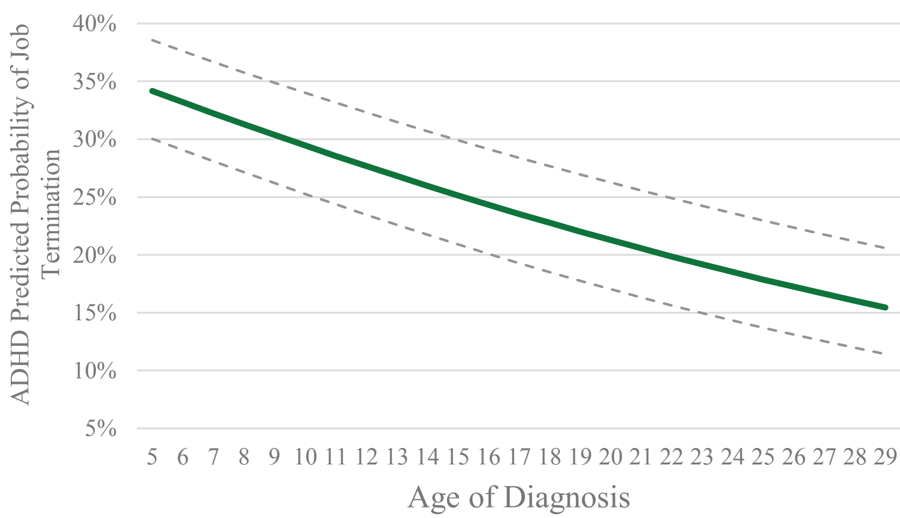
Regressors	Model 1 Age of Diagnosis > 0	Model 2 Age of Diagnosis ≥ 5
Diagnosed with ADHD	0.2433**	0.3016**
	(0.0999)	(0.1182)
Age of ADHD Diagnosis	−0.0077*	−0.0112**
	(0.0042)	(0.0052)
HH Income	−0.0012**	−0.0013**
	(0.0006)	(0.0006)
Female	−0.1214***	−0.1158***
	(0.0447)	(0.0424)
Black	0.1193***	0.1173***
	(0.0456)	(0.0445)
Hispanic	-0.0199	-0.0170
	(0.0203)	(0.0192)
High School Marijuana	0.0321*	0.0302*
	(0.0171)	(0.0163)
High School Alcohol	0.0324*	0.0314**
	(0.0166)	(0.0160)
Akaike Information Criterion	12955	12917
n	11,102	11,083

Heteroskedasticity-robust standard errors reported in parentheses
Post stratified untrimmed cross-sectional grand sample weight used to estimate models.
**** implies estimate is significant at 1%*
*** implies estimate is significant at 5%*
** implies estimate is significant at 10%*

When deriving the sample size for the job termination model, the same factors that influenced the earnings models' final sample hold here also. In this case, however, there were only 516 unusable responses from the dependent variable which increased the sample size overall.

Like the results of the earnings models, the results of the job termination models are difficult to interpret fully from Table 3. ADHD's effect on the probability of being fired is captured by two variables and, therefore, must be interpreted simultaneously. Figure 7 illustrates the results from Model 2.

Figure 7: ADHD's Predicted Impact on Job Termination



The solid line is ADHD's effect on the predicted probability of job termination. The dashed lines represent a 95% confidence interval for the prediction.

Based on Hypothesis 4, the results from the job termination model yield exactly the opposite outcome from what was anticipated. Rather than a relatively early diagnosis' reducing the probability of being fired, these results argue that an early diagnosis *increases* the ADHD agent's chances of being fired. Similarly to the earnings model, the slope of the predicted function in Figure 7 is the effect of the age of diagnosis. Since both estimates are significant and negative in Models 1 and 2, the effect of the age of diagnosis is robust in the opposite direction of expectations.[22] While I anticipated the age of diagnosis to have a positive relationship with the probability of being fired, it has a negative relationship in these models.

This creates an interesting tension with the results from the earnings model. While being diagnosed from ages 5 to 14 arguably yields strong benefits in regard to earnings,

[22]The slope of Figure 7 is the partial derivative of equation 25 with respect to the age of diagnosis which is simply the estimated coefficient =-0.0112 for Model 2

the ADHD agent who is diagnosed this early in life is still more vulnerable to job termination than someone diagnosed later in life. Since job termination arguably incurs serious short-term and potentially long-term costs to the agent, the benefits and costs of an ADHD diagnosis need to be weighed against each other carefully.[23] These results are discussed further in Section 5.

5 Discussion of Results

For the earnings models in Table 2, Models 1 and 3 argue not to reject Hypothesis 1 since the estimated parameter is negative and significant. When the age of diagnosis is restricted to a minimum of 5 years old, however, Models 3 and 4 argue to reject Hypothesis 1 since the estimated parameter is no longer significant. When the age of diagnosis is incorporated, the results in Figure 6 offer evidence that an agent who is diagnosed from ages 5 to 14 can entirely offset this negative effect of having ADHD in the labor market; therefore, I do not reject Hypothesis 3.

The results of the job termination model argue not to reject Hypothesis 2 since ADHD increases the chances of experiencing job termination in both Models 1 and 2 of Table 3. When the age of diagnosis is included, Figure 7 shows the interesting result that an earlier age of diagnosis leads to a higher probability of job termination; therefore, I reject Hypothesis 4. This result is not entirely unexpected given the theoretical discussion provided in Section 2, but it creates an interesting contrast when paired with the results of the earnings model.

It remains to give economic justification to the puzzle between the earnings and job

[23] For instance, having a break in employment due to job termination might prevent him from receiving gainful employment in the future due to the stigma of being fired.

termination outcomes. Why might someone diagnosed relatively early in life go on to receive the same earnings yet have a higher chance of being fired compared to someone diagnosed later in life? And how might either of the theoretical models presented in Section 2 help explain these results?

Focusing on the later end of Figure 7 where the age of diagnosis is relatively high, how might this puzzle be explained? Consider the issue of endogeneity. If a non-ADHD agent has been in the labor market for several years yet is not earning as much as he would like, he might be incentivized to get tested for ADHD to explain his lack of productivity. Considering the ambiguity with which ADHD is diagnosed, the probability that this agent is diagnosed might be fairly high.[24] He has been in the labor market so long, however, that he understands how to survive. Thus, both his earnings are low and his probability of being fired is low.

Now, looking to the front end of Figure 7 where the age of diagnosis is relatively young, one might consider the time preference of the agent. If the agent is highly patient, he will select into a job more honestly than an agent who is highly impatient. That is, a patient ADHD agent will select into a job where he maximizes his long-run benefits whereas an impatient or myopic ADHD agent selects a job to maximize his short-run benefits.

For the patient agent, this might mean selecting a particular occupation where he receives lower earnings in the short run but can be most productive in the long run. It might be safe to assume that a non-ADHD agent is better suited for high-skill, low-risk work than a non-ADHD agent, and that an ADHD agent is better suited for low-skill,

---

[24] The "ADHD as a social construct" perspective would tend to follow this line of thinking (Timimi 2004). That is, this relatively new phenomenon of ADHD might simply be a scapegoat for a person's low productivity caused by something entirely different such as poor work environment or personal distractions.

high-risk work than a non-ADHD agent.[25] By selecting into this different line of work, an ADHD agent can earn the same as a non-ADHD agent. Considering he is involved in an entirely different echelon of work, however, the nature of the occupation might lend itself to termination more so than the non-ADHD agent's occupation. That is, the patient agent is more exposed to macroeconomic labor market forces influencing job terminations rather than being fired because of his lack of performance.

For the impatient agent, this might mean selecting the same high-skill job as a non-ADHD agent so he earns as much as the non-ADHD agent in the short run but might be less productive in the long run. From here, the impatient agent leaves himself vulnerable to job termination since the principal will have entered a suboptimal contract with him.

In either case, these explanations might serve as suitable economic justifications since the ADHD agent earns the same amount as a non-ADHD agent yet is still susceptible to job termination more than the non-ADHD agent. The theoretical models will have to be developed with future empirical results to support or reject an explanation to the puzzle.

Regardless of the explanation, these results show how important information is in various markets. The health economics related to ADHD might easily influence the traditional labor economics of the disorder. For instance, if physicians are incentivized to diagnose ADHD and prescribe medication to treat it because of contracts with pharmaceutical companies or expanded access to health insurance (Currie, Stabile, Jones 2014), the transaction taking place in that healthcare market might have a significant impact on the agent's future transactions and general well-being in the labor market.

[25]High-risk jobs might include mining, oil rigging, or construction work, and they require little formal education and maybe some more specialized technical training. Yet these jobs pay wages above competitive levels since the agent must be compensated for the high risk he bears.

Furthermore, if public funds tied to school or teacher performance influence the diagnosis of ADHD, then these transactions in the public school market might significantly impact the agent's outcome in the labor market. Incentivizing as efficient or perfect information as possible when diagnosing ADHD is especially pertinent since the information derived from a diagnosis strongly influences how the agent makes his decisions and how others perceive him (Hinshaw, Stier 2008).[26] As long as a disorder like ADHD is susceptible to imperfect information, misdiagnosing people due to misaligned incentives can carry consequences for them in other facets of life.

5.1 Future Research

The puzzle illustrated in section 4 offers much more to be explored in the economics of ADHD and mental disorders in general.

Firstly, regarding the issue of endogeneity as one possible explanation, there are statistical considerations that ought to be taken into account given the data. Figure 9 in the appendix is one example. Figure 9 shows that as the age in Wave IV increases, the age of diagnosis also tends to increase. There might be a trend here or even issues of reverse causality that should be accounted for in future estimations.[27]

Using a strong instrument variable for an ADHD diagnosis would be an interesting way to control for potential reverse causation.

Otherwise, improving the data collection regarding the diagnosis of ADHD can also prevent any endogeneity in the models. For such a nuanced disorder in as nuanced a discipline as economics, there are many more considerations that need to be accounted

[26] Hinshaw, S., Andrea Stier. 2008. "Stigma as Related to Mental Disorders." Annual Review of Clinical Psychology. 4:367-393.

[27] Refer to Section 3.2.4 for an explanation of how reverse causality was controlled to an extent in this thesis.

for but which Add Health simply does not control for in its surveys. Being able to obtain a more reliable report of an ADHD diagnosis is imperative when minimizing measurement error in the models.

Secondly, the results in this thesis argue that it is important to understand how ADHD agents select into jobs, particularly as it pertains to the age of diagnosis. If people diagnosed relatively early select occupations differently than those diagnosed late, this might reveal more about agents' behavior and a clearer reality of the diagnosis of the disorder. For example, if an early diagnosis has a higher probability of being a true diagnosis (i.e. not a "social construct" diagnosis (Timimi 2004)), then those with a relatively later diagnosis might not actually be ADHD agents. Furthermore, those with an early diagnosis might be myopic and select into the high-paying jobs, increasing their chances of job termination. In general, testing to determine if true ADHD agents tend to be myopic could help inform the literature.

Other modifications and future research questions are also generated from these results. For example, it might be interesting to test a threshold model for evidence of a threshold age of diagnosis at which any benefits of early diagnosis do not outweigh the costs. This might lead to interesting policy proposals regarding the minimum age of diagnosis.

Considering Add Health provides a count variable for the age of diagnosis, estimating a Poisson regression model for the number of job terminations that ADHD agents experience compared to non-ADHD agents would further contribute to our understand of ADHD's impact on job termination.

Another specification adjustment might be to redefine the age variable in the labor market or adjust how ADHD is measure in the model. For instance, the models in this thesis simply use the agent's age-level at Wave IV and the age of diagnosis if it applies. By taking the difference between the Wave IV age and the age of diagnosis, this might be a more effective way of capturing the effect of the age of diagnosis. By measuring the effect of ADHD this way, it might reveal more about how people adjust to new information about their own productivity before they enter the labor market.

Besides modified model specifications, it might be useful to select different samples for the models to see how robust the results are. For example, by running the same models in this thesis except with only the ADHD sample (about 775 observations) might offer a clearer picture of exactly how the age of diagnosis influences outcomes.

Finally, further study of the economics of ADHD ought to control for treatment. The two-stage human capital investment model in Section 2.3 offers a simple foundation from which to build further research of the disorder's treatment. Since the psychostimulants often used to treat ADHD directly impact people's productivity, there are myriad economic questions available to study. For instance, if the drugs are meant to bring ADHD agents up to par with non-ADHD agents' productivity, how might barriers to nonmedical access influence the overall economic effect of the drugs? If barriers to access are low enough for non-ADHD agents to use the drugs to boost their "normal" productivity even higher, there might be strong gains-from-trade for doing so. The non-ADHD agent will take advantage of the opportunity to use his time more freely if he can work more quickly and efficiently with the help of the drugs. From the ADHD agent's perspective, though, the productivity gap is now the same as before. His productivity has increased under the drugs but so has the non-ADHD agent's productivity.

All of these future lines of research can significantly contribute to our understanding of the ADHD conundrum. Developing a clearer economic understanding of ADHD can provide more perfect information than the medical or psychology profession alone can offer. Moving toward more perfect information will improve households' and even physicians' decision-making regarding ADHD, which can improve social welfare in general.

References

Acemoglu, D. and J. Angrist (1998). Consequences of employment protection? the case of the americans with disabilities act. Technical report, National bureau of economic research.

Acemoglu, D. and D. Autor (2009). Lectures in labor economics. *Unpublished manuscript, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA. Retrieved from http://economics. mit. edu/files/4689*.

Association, A. P. et al. (2013). *Diagnostic and statistical manual of mental disorders, (DSM-5®)*. American Psychiatric Pub.

Baldwin, M. L. (1999). The effects of impairments on employment and wages: Estimates from the 1984 and 1990 sipp. *Behavioral sciences & the law 17*(1), 7–27.

Baldwin, M. L. and S. C. Marcus (2007). Labor market outcomes of persons with mental disorders. *Industrial Relations: A Journal of Economy and Society 46*(3), 481–510.

Becker, G. S. and C. B. Mulligan (1997). The endogenous determination of time preference. *The Quarterly Journal of Economics*, 729–758.

Bloom, B., R. A. Cohen, and G. Freeman (2010). Summary health statistics for us children: National health interview survey, 2009. *Vital and health statistics. Series 10, Data from the National Health Survey* (247), 1–82.

Bokhari, F. A. and H. Schneider (2011). School accountability laws and the consumption of psychostimulants. *Journal of Health Economics 30*(2), 355–372.

Breining, S. N. (2014). The presence of adhd: Spillovers between siblings. *Economics Letters 124*(3), 469–473.

Cascade, E., A. H. Kalali, and S. B. Wigal (2010). Real-world data on: attention deficit hyperactivity disorder medication side effects. *Psychiatry (Edgmont) 7*(4), 13.

Currie, J. and M. Stabile (2006). Child mental health and human capital accumulation: the case of adhd. *Journal of health economics 25*(6), 1094–1118.

Currie, J., M. Stabile, and L. Jones (2014). Do stimulant medications improve educational and behavioral outcomes for children with adhd? *Journal of health economics 37*, 58–69.

DeLeire, T. (2000). The wage and employment effects of the americans with disabilities act. *Journal of Human Resources*, 693–715.

DeLeire, T. (2001). Changes in wage discrimination against people with disabilities: 1984-93. *Journal of Human Resources*, 144–158.

Doshi, J. A., P. Hodgkins, J. Kahle, V. Sikirica, M. J. Cangelosi, J. Setyawan, M. H. Erder, and P. J. Neumann (2012). Economic impact of childhood and adult attention-deficit/hyperactivity disorder in the united states. *Journal of the American Academy of Child & Adolescent Psychiatry 51*(10), 990–1002.

Fletcher, J. and B. Wolfe (2008). Child mental health and human capital accumulation: the case of adhd revisited. *Journal of health economics 27*(3), 794–800.

Fletcher, J. M. (2014). The effects of childhood adhd on adult labor market outcomes. *Health economics 23*(2), 159–181.

Getahun, D., S. J. Jacobsen, M. J. Fassett, W. Chen, K. Demissie, and G. G. Rhoads (2013). Recent trends in childhood attention-deficit/hyperactivity disorder. *JAMA pediatrics 167*(3), 282–288.

Goldman, L. S., M. Genel, R. J. Bezman, P. J. Slanetz, et al. (1998). Diagnosis and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Jama 279*(14), 1100–1107.

Harris, K. M. (2009). The national longitudinal study of adolescent to adult health (add health), waves i ii, 1994-1996; wave iii, 2001-2002; wave iv, 2007-2009 [machine-readable data file and documentation]. Carolina Population Center, University of North Carolina at Chapel Hill.

Hinshaw, S. P. and R. M. Scheffler (2014). *The ADHD Explosion: Myths, Medication, Money, and Today's Push for Performance.* Oxford University Press.

Hinshaw, S. P. and A. Stier (2008). Stigma as related to mental disorders. *Annu. Rev. Clin. Psychol. 4*, 367–393.

Langberg, J. M. and S. P. Becker (2012). Does long-term medication use improve the academic outcomes of youth with attention-deficit/hyperactivity disorder? *Clinical child and family psychology review 15*(3), 215–233.

Matza, L. S., C. Paramore, and M. Prasad (2005). A review of the economic burden of adhd. *Cost effectiveness and resource allocation 3*(1), 1–9.

Mincer, J. (1974). Schooling, experience, and earnings. human behavior & social institutions no. 2.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

Polanczyk, G., M. S. de Lima, B. L. Horta, J. Biederman, and L. A. Rohde (2007). The worldwide prevalence of adhd: a systematic review and metaregression analysis. *The American journal of psychiatry 164*(6), 942–948.

Scheffler, R. M., T. T. Brown, B. D. Fulton, S. P. Hinshaw, P. Levine, and S. Stone (2009). Positive association between attention-deficit/hyperactivity disorder medication use and academic achievement during elementary school. *Pediatrics 123*(5), 1273–1279.

Schneider, H. and D. Eisenberg (2006). Who receives a diagnosis of attention-deficit/hyperactivity disorder in the united states elementary school population? *Pediatrics 117*(4), e601–e609.

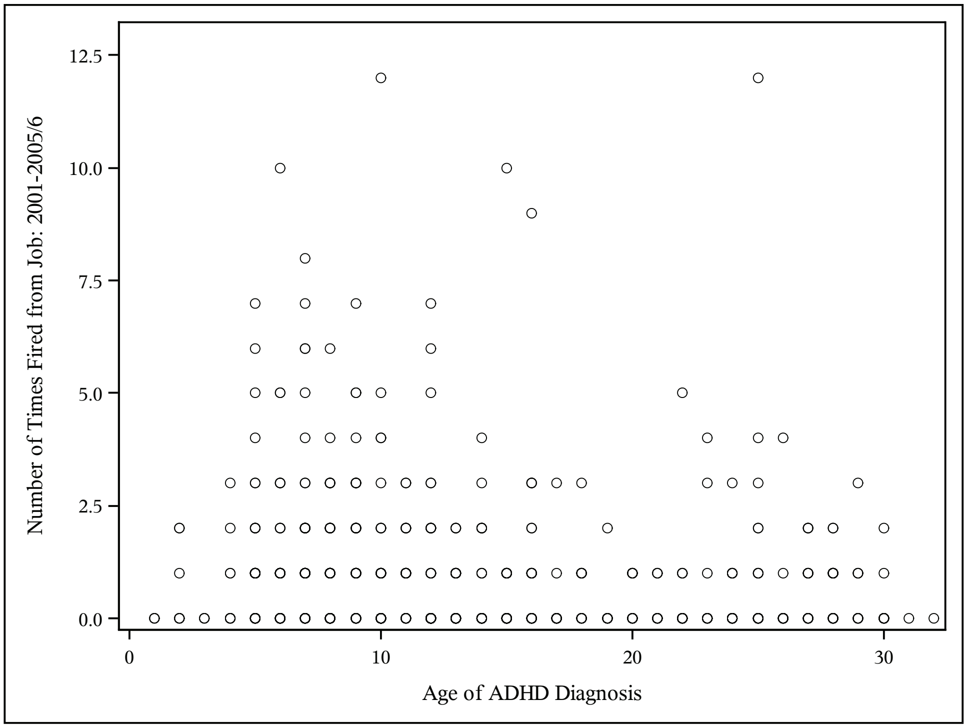
Timimi, S. and E. Taylor (2004). Adhd is best understood as a cultural construct. *The British Journal of Psychiatry 184*(1), 8–9.

Wehmeier, P. M., A. Schacht, and R. A. Barkley (2010). Social and emotional impairment in children and adolescents with adhd and the impact on quality of life. *Journal of Adolescent Health 46*(3), 209–217.

Yelin, E. H. and M. G. Cisternas (1997). Employment patterns among persons with and without mental conditions. *Mental disorder, work disability and the law*, 25–51.

6 Appendix

Figure 8: Age of Diagnosis and Wave IV Job Terminations



Also like the graphs of earnings and age of diagnosis, there appears to be somewhat of a cubic relationship here. This relationship is less intuitive than the earnings one, however, since we would anticipate fewer job terminations if the agent is diagnosed earlier in life. Figure 8 shows that an early diagnosis can maximize his number of firings which is counterintuitive to the hypothesis.

6.1 Full Earnings Model Results

Table 4: Results of Earnings Models

	Mincer Models		Level Models	
	Model 1	Model 2	Model 3	Model 4
Regressors	Age of Diagnosis > 0	Age of Diagnosis ≥ 5	Age of Diagnosis > 0	Age of Diagnosis ≥ 5
Intercept	4.4767	4.4572	-203982	-198698
	(4.2293)	(4.2459)	(131898)	(133121)
Some High School	0.3664*	0.3633*	12430***	12480***
	(0.2193)	(0.2203)	(4443.94)	(4431.36)
High School	0.5406**	0.5330**	12419***	12475***
	(0.2128)	(0.2138)	(3092.41)	(3086.73)
Technical Training	0.7910***	0.7843***	16292***	16333***
	(0.2095)	(0.2105)	(3060.78)	(3053.13)
Some College	0.7772***	0.7732***	17620***	17683***
	(0.2082)	(0.2091)	(2943.62)	(2931.34)
Bachelor's Degree	1.1176***	1.1126***	30156***	30217***
	(0.2095)	(0.2104)	(3152.46)	(3142.61)
Master's Degree	1.1253***	1.1191***	28229***	28273***
	(0.2137)	(0.2147)	(3239.00)	(3230.89)
Doctoral Degree	1.3753***	1.3709***	43617***	43672***
	(0.2201)	(0.2210)	(4948.73)	(4940.47)
Diagnosed with ADHD	−1.3695**	-1.7511	−26462**	-21618
	(0.6716)	(1.0902)	(12138)	(18735)
Age of ADHD Diagnosis	0.2772*	0.3603	5645.54*	4599.18
	(0.1528)	(0.2382)	(3008.17)	(4312.36)
Age of ADHD Diagnosis^2	−0.0196*	-0.0249	−407.36*	-341.23
	(0.0105)	(0.0155)	(209.69)	(285.61)
Age of ADHD Diagnosis^3	0.0003954*	0.0004962	8.1410*	6.8818
	(0.0002118)	(0.0003020)	(4.2950)	(5.6413)
Age	0.2672	0.2689	13620	13244
	(0.2872)	(0.2884)	(9065.45)	(9153.11)
Age^2	-0.0039	-0.0039	-201.61	-195.00
	(0.0049)	(0.0050)	(154.82)	(156.36)
ln(HH Income)	0.1470***	0.1474***		
	(0.0296)	(0.0296)		
HH Income			54.28***	54.23***
			(16.3197)	(16.3264)
Northeast	0.1232**	0.1249**	4067.06**	4091.34**
	(0.0595)	(0.0595)	(2004.03)	(2004.97)
Midwest	-0.0089	-0.0078	−3207.25***	−3223.25***
	(0.0573)	(0.0573)	(1234.97)	(1235.91)
South	0.0409	0.0404	-146.45	-162.46
	(0.0576)	(0.0576)	(1385.80)	(1387.76)
Female	−0.4079***	−0.4088***	−12640***	−12662***
	(0.0318)	(0.0319)	(1056.04)	(1056.71)
Black	−0.1321***	−0.1346***	−4957.63***	−4925.07***
	(0.0429)	(0.04296)	(1357.70)	(1362.68)
Hispanic	0.0521	0.0526	-9.87	-12.07
	(0.0616)	(0.0616)	(1432.85)	(1433.20)
High School Marijuana	−0.0656*	−0.0661*	-1048.81	-1047.34
	(0.0345)	(0.0345)	(1325.72)	(1326.46)
High School Alcohol	0.1122***	0.1122***	3086.57***	3087.89***
	(0.0329)	(0.0328)	(1129.79)	(1129.92)
Adjusted R-squared	0.1363	0.1362	0.0755	0.0753
n	9,850	9,837	9,850	9,837

Heteroskedasticity-robust standard errors reported in parentheses
Post stratified untrimmed cross-sectional grand sample weight used to estimate models.
**** $p < 0.01$*
*** $p < 0.05$*
** $p < 0.10$*

6.2 Full Job Termination Model Results

Table 5: Results of Job Termination Model

Regressors	Age of Diagnosis > 0	Age of Diagnosis ≥ 5
Intercept	-0.3680**	-0.3555**
	(0.1498)	(0.1439)
Some High School	0.3139**	0.3035**
	(0.1343)	(0.1292)
High School	0.2352**	0.2311**
	(0.1117)	(0.1082)
Technical Training	0.2405**	0.2356**
	(0.1136)	(0.1100)
Some College	0.1993*	0.1955**
	(0.1019)	(0.0986)
Bachelor's Degree	0.0961	0.0965
	(0.0815)	(0.0787)
Master's Degree	-0.0801	-0.0708
	(0.0861)	(0.0818)
Doctoral Degree	-0.2433*	-0.2294*
	(0.1395)	(0.1330)
Diagnosed with ADHD	0.2433**	0.3016**
	(0.0999)	(0.1182)
Age of ADHD Diagnosis	-0.0077*	-0.0112**
	(0.0042)	(0.0052)
HH Income	-0.0012**	-0.0013**
	(0.0006)	(0006)
Northeast	0.0524*	0.0528*
	(0.0286)	(0.0279)
Midwest	0.0435*	0.0428*
	(0.0236)	(0.0229)
South	-0.0189	-0.0176
	(0.0189)	(0.0181)
Female	-0.1214***	-0.1158***
	(0.0447)	(0.0424)
Black	0.1193***	0.1173***
	(0.0456)	(0.0445)
Hispanic	-0.0199	-0.0170
	(0.0203)	(0.0192)
High School Marijuana	0.0321*	0.0302*
	(0.0171)	(0.0163)
High School Alcohol	0.0324*	0.0314**
	(0.0166)	(0.0160)
Akaike Information Criterion	12955	12917
n	11,102	11,083

Heteroskedasticity-robust standard errors reported in parentheses
Post stratified untrimmed cross-sectional grand sample weight used to estimate models.
**** $p < 0.01$*
*** $p < 0.05$*
** $p < 0.10$*

Figure 9: Age vs. Age of Diagnosis