School accountability laws and the consumption of psychostimulants [†]

Farasat A.S. Bokhari [a,b,*], Helen Schneider [c]

[a] *Division of Health and Social Care Research, Kings College London, London SE1 3QD*

[b] *LSE Health and Social Care, London School of Economics and Political Science, London WC2A 2AE*

[c] *Department of Economics, University of Texas at Austin, TX 78712*

March 8, 2009
(Revised: December 14, 2010)

Abstract

Over the past decade, several states introduced varying degrees of accountability systems for schools, which became federal law with the passage of the No Child Left Behind Act of 2001. The intent of these accountability laws was to improve academic performance and to make school quality more observable. Nonetheless, schools have reacted to these pressures in several different ways, some of which were not intended. We make use of the variation across states and over time in specific provisions of these accountability laws and find that accountability pressures effect medical diagnoses and subsequent treatment options of school aged children. Specifically, children in states with more stringent accountability laws are more likely to be diagnosed with Attention Deficit/Hyperactivity Disorder (ADHD) and consequently prescribed psychostimulant drugs for controlling the symptoms. However, conditional on diagnosis, accountability laws do not further change the probability of receiving medication therapy.

Key words: Attention Deficit Hyperactivity Disorder, ADD/ADHD, psychostimulants, school accountability laws

JEL Classification: I12, I28, H75

Published:

Journal of Health Economics 30(2), 355-372, 2011

[†] This study was funded in part by the Economics Department at Florida State University that helped purchase the data sets. Thanks to Krista Jabs for her assistance in putting together some of the data sets and to Nadine Zubair for her support. Special thanks to Daniel Eisenberg, Tim Sass, and David Figlio for their comments and feedback on an an earlier draft. We are also indebt to the detailed comments provided by anonymous referees. The usual disclaimer applies.

[*] Corresponding Author: Farasat Bokhari (+44 (0) 207-848-8229)

Email addresses: farasat.bokhari@kcl.ac.uk (Farasat A.S. Bokhari), hs3473@eco.utexas.edu (Helen Schneider).

1. Introduction

Over the past decade, state level accountability systems were an important part of the education reform that aimed to unify state proficiency requirements in subjects like math and reading. Public reporting of school performance on standardized tests had an additional benefit of achieving some transparency of school quality. The presence of accountability laws varied widely from state to state, and over time until the passage of the No Child Left Behind Act (NCLB) of 2001, which required states to create an accountability system of assessments, graduation rates, and other indicators. The Act mandates that states administer high quality annual assessments to every child from grades three through eight, and must be aligned to standards consistent with nationally recognized professional and technical standards (U.S. Department of Education, 2002). The accountability provisions of the NCLB require schools to make Adequate Yearly Progress (AYP), based on student performance on these standardized tests.

The response of schools to accountability pressures is widely debated among researchers, specifically the means that schools can employ to increase the number of students who achieve proficiency on mandatory standardized tests in order to avoid sanctions if they fail to meet state targets. In an earlier review of the literature, (Koretz, 2002) found that the apparent gains in student academic performance may be, in some cases, illusory. For example accountability pressures generate perverse incentives to inflate scores, especially on high stakes tests that affect the school (Koretz, 2002). Specifically, studies have found sharp gains on high stakes tests accompanied by no gains on audit tests (Koretz, 1988, Koretz and Barron, 1998). In more recent studies, Neal and Schanzenbach (2007) and Reback (2008) using two separate data sets from Chicago and Texas respectively, document that schools have an incentive to improve the academic performance of students who are on the margin of passing since it is the passing rates that are being reported rather than the absolute scores. Similarly, Figlio (2006) finds that in response to accountability pressures, schools in Florida re-shaped the testing pool through selective disciplining.

In this paper we show that school accountability laws also effect medical diagnoses and subsequent treatment options: children in states with more stringent accountability laws are more likely to be diagnosed with Attention Deficit/Hyperactivity Disorder (ADHD), and consequently, prescribed psychostimulant drugs for controlling the symptoms.[1] ADHD is a psychiatric condition which has an estimated prevalence of nearly 8% in school aged children in the U.S., and about 60% of these children are prescribed medication for the disorder (Centers for Disease Control and Prevention, 2005). Our analysis has uncovered an unintended consequence of public policy that should be part

[1]Psychostimulants are the most commonly prescribed drugs for the treatment of ADHD, and include Schedule II drugs such as Methyplhenidates (e.g Ritalin, Concerta), Mixed Amphatamine Salts (e.g Adderall) and Dextroamphatamines (e.g Dexedrine). Other molecules used for treatment of ADHD include Atamoxatine (Stratera) which is not a stimulant drug.

of the debate over school reform but also exposes that medical decision making itself is also affected by seemingly unrelated policies. While this is not the first paper to show unintended consequences of accountability pressures, it does provide a direct evidence of the extent of the spill-over into a child's health.[2]

Following the passage of the NCLB, all states implemented the testing (since NCLB requires reporting of Adequate Yearly Progress by schools) and most states made report cards and ratings of schools publicly available. However, many states continued or created a system of rewards and sanctions based on whether or not students from all the different subgroups in their schools make adequate yearly progress, which is measured by how many students pass the state standardized tests. Thus, although the NCLB requires a certain minimum level of accountability to be ensured by the states, accountability standards such as the presence of school assistance, rewards and sanctions vary from state to state. Each state determines appropriate levels of proficiency for its students and can choose to reward or sanction schools based on their performance.

We show that accountability pressures affect the probability of diagnosis and medication for ADHD. We exploit the variation in specific provisions of state laws regarding accountability that vary by state and years and, link this information to two unique data sets on medication and diagnosis. The first data set provides aggregate measures of the consumption of all psychostimulant drugs by state and year between 1999 and 2003, while the second data set is nationally representative data of all school aged children for 2003, and includes information on individual and family characteristics as well as the children's health status, including if they are ADHD and on medication for it.

With the first data set, we use a difference-in-difference model and take advantage of both, between-state variation in accountability laws, as well as variation over time to show how changes in school pressures to improve academic performance effect overall psychostimulant consumption. In the second data set, we supplement our analysis of the U.S. per-capita consumption level with an analysis of individual, student-level diagnoses of ADHD and psychostimulant consumption for children enrolled in public schools. This individual micro-level data allows us to control for child and family characteristics to estimate the marginal impact of accountability laws on the probability of ADHD diagnosis and medication therapy. Additionally, using the micro-level data we construct a falsification test for the effect of these laws: it is possible that accountability laws, which are at the state level, are correlated with other unobserved state level factors that are responsible for diagnosis of ADHD (and not the laws per se). Hence, we repeat the analysis on children enrolled in private

[2] Anderson and Butcher (2006) show that new accountability measures along with other factors such as population growth and property-tax restrictions pressure schools to raise additional funds through vending contracts and snack food sales. Such pressures translate into unhealthy school food policies and contribute to childhood obesity. Figlio and Winicki (2005) found that schools in Virginia respond to accountability pressures by offering lunches with higher caloric intake on test days to improve standardized test scores. Some other studies have also shown that schools tend to respond by classifying more marginal students as disabled (Figlio and Getzler, 2006, Cullen and Reback, 2006, Jacob, 2005).

schools, a group that is not subject to the accountability laws, but have otherwise similar risk factors associated with being diagnosed with ADHD. For this subgroup we find that the state laws do not change the probability of diagnosis, lending credence to our claim that the accountability laws are not just reflecting the effect of other state level unobserved factors that effect the diagnosis of ADHD.

Our results indicate that provisions of assistance and rewards are associated with a statistically significant increase in aggregate psychostimulant consumption (3.1% and 2.6% respectively) while sanctions on schools are not significant. Alternative school accountability indexes that measure the presence of 'consequential' accountability or the overall strictness of the accountability laws also show similar associations (elasticity with respect to strictness index ranges from .027 to .031 and is statistically significant). Further, these provisions increase the probability of being diagnosed with ADHD for an individual school aged child by .018 and .012 for assistance and rewards respectively, while the provision for imposing sanctions on poor performing schools is not significant. Similarly, the marginal effect of the accountability indexes (strictness measures) are also positive and significant. Finally, the accountability laws do not change the probability of receiving medication therapy conditional on being diagnosed with ADHD. Thus, medication therapy and overall consumption of stimulants increases because more children are diagnosed with ADHD rather than because more children who are ADHD are prescribed medication therapy. According to our simulations, about 1.1 million *fewer* public school children would be diagnosed with ADHD when no accountability laws are present versus when all three laws are present.

It is important to note that the potential increase in ADHD diagnosis and psychostimulant consumption rates that follows a hike in school accountability pressures is not necessarily undesirable. On one hand, some schools may be "gaming the system" by inappropriately labeling marginal students as ADHD to provide them with accommodations as well as to reshape the testing pool. On the other hand, stricter accountability leads to more ADHD students receiving the appropriate diagnosis, academic accommodations and efficacious medical treatment that improves their academic performance, as well as of their peers. While physician decisions on ADHD cases are certainly influenced by input from school personnel (as we later argue), our study provides a possible opportunity for interpreting whether the behavioral consequences of accountability have been positive or negative along this one dimension. This can be important because many children with ADHD, if untreated, continue to exhibit symptoms of the disorder into adulthood, where such symptoms impair activities of daily living, educational achievement and productivity (Murphy and Barkley, 1996, Secnik et al., 2005, Biederman and Faraone, 2006). For example, Biederman and Faraone (2006) found that subjects in the ADHD group are less likely to pursue education beyond some high school or to hold a full time job, and computed loss of workforce productivity associated with

ADHD between $67 billion and $116 billion. Therefore, early identification and treatment may alleviate some of these workforce losses.

The link between accountability pressures and ADHD diagnosis and medication therapy requires both the incentives as well as the ability, of school personnel to influence medical decision making. In the next section, we argue that schools do have both the incentive and the ability to influence the diagnosis of ADHD. Section three describes the provisions of the state level accountability laws that we use in our analysis and provides details of our data. The fourth section provides descriptive statistics as well as the main results. This is followed by a brief summary and conclusions section.

2. Role of School Personnel in ADHD Diagnosis

Schools may have strong incentives to label a child as ADHD for several reasons. First, ADHD is a disorder associated with significant impairments that commonly continue into adulthood, including poorer performance and earlier exit from school (Mannuzza et al., 1997). Once a child is diagnosed, s/he may receive psychostimulant and other treatments. The psychostimulant treatments have been established to be efficacious, resulting in improved classroom behavior as well as some improvements in academic achievement: Carlson et al. (1992) found that both methylphenidates and behavior modification alone significantly improved children's classroom behavior, but only methylphenidates improved children's academic productivity and accuracy. Evans and Pelham (1991) in a double-blind, placebo-controlled assessment found significant effects of psychostimulants on quiz and test performance, observations of attention and behavior during lectures, teacher ratings, as well as accuracy on assignments completed during study hall. Similarly, A Multimodal Treatment Study of Children with ADHD used a randomized treatment design and followed a cohort of 579 ADHD children over 14 months. The study found a significant reduction in symptoms over time and children who received psychostimulant treatment showed significantly greater improvement than those given intensive behavioral treatment and community care (MTA Cooperative Group et al., 1999a,b). More recently, a comprehensive review of the current literature by The American Academy of Pediatrics (Brown et al., 2005) found that current empirical evidence strongly supports the use of stimulant medications for treating the core symptoms of children with ADHD with significant effects on measures of attention, distractibility, and impulsivity and observable social and classroom behavior. Some modest effects were also found for academic achievement with effect sizes of 0.19-0.47 with mean of 0.34. Finally, Barbaresi et al. (2006) in a population-based study found that on average psychostimulant treatment was modestly correlated with improved reading achievement scores (ρ = .15, p = .012), that both treatment with psychostimulants and longer duration of medication were associated with decreased absenteeism, and that children with ADHD who were treated with stimulants were 1.8 times less likely to be retained a grade than children with ADHD who were not treated (however, they found no association between psychostimulant treatment and school dropout rates). Since ADHD diagnosis leads to psychostimulant treatment for 60% of affected children (Centers for Disease Control and Prevention, 2005), schools facing stricter accountability laws have an incentive to assist parents and physicians in both ADHD identification and treatment. Such treatment has been shown to enhance academic performance; for instance, both reading and math performance tested under NCLB, are seen to exhibit sizeable improvement (Scheffler et al., 2009).

Second, in addition to the improvements in behavior and academic achievement of the child with ADHD, treatment may also improve the achievement of other children in the classroom via peer effects. Lazear (2001) develops an education production model and shows that the ability of a student to learn depends on the behavior of his/her classmates because disruptive behaviors of peers reduce effective teaching time and may directly interfere with school work. Thus, if an ADHD student receives the appropriate diagnosis and treatment, academic performance of the entire class may improve. Empirical evidence supports the presence of negative externalities of ADHD in classroom settings. Aizer (2008) reports that children with undiagnosed ADHD lower the reading test scores of non-ADHD classmates (if 8.5 percent of the class have undiagnosed ADHD, test scores will be 2 points or 20 percent of a standard deviation, lower) but once these children are diagnosed, no such externalities are observed. Moreover, ADHD students put more stress on teachers as measured by Index of Teaching Stress (Greene et al., 2002) and psychostimulant treatment has been shown to improve teacher's ratings of children's ADHD symptoms (Pelham et al., 2000).

Third, ADHD is a disability that is recognized by the Individuals with Disabilities in Education Act (IDEA). Thus, all ADHD students are entitled to academic accommodations, although accommodations under IDEA vary across states.[3] Additionally, IDEA requires states to include students with disabilities in their testing systems and report the results. Under the NCLB Act schools have to test 95% of all students, including those with disabilities (Ansell, 2004). Some states report proficiency rates on state standardized tests for students with learning disabilities separately since many of them receive accommodations.[4] Identifying a previously undiagnosed child with ADHD could improve the academic performance of students in both pools, students with learning disabilities (since, as previously noted, psychostimulant treatments have been established to be efficacious for children with ADHD) as well as for the general student population since ADHD kids do worse than typical

[3] Under the policy change to IDEA in 1990, the Department of Education issued a policy clarification memorandum in 1991 stating that schools not only had to provide special services and accommodations for children with sufficiently severe ADHD, they had to evaluate all children suspected by their parents and local education agencies of having the disorder (Aleman, 1991). Currently, the U.S. Department of Education states that children diagnosed with ADHD are eligible for special education services and are categorized under "Other Health Impaired" (Sec. 300.7) group of disabilities. Also, the new regulations implementing the IDEA Amendments of 1997 (issued March 12, 1999) for the first time explicitly incorporate ADHD within the definition of "Other Health Impaired".

[4] For instance, the Adequate Yearly Progress (AYP) requirements of NCLB also require schools to meet benchmarks for distinct sub-populations, one of which is students with disabilities.

peers. Therefore, ensuring diagnosis and psychostimulant treatment to ADHD students serves as an effective strategy that schools and teachers may employ to meet state proficiency requirements.

Although schools and teachers are not physicians and cannot diagnose or prescribe drugs, they have a strong influence on the medical decision making process. Since there are no laboratory tests that can be performed to diagnose ADHD and other such learning problems, the diagnoses of ADHD have always been controversial. To meet the diagnostic criteria for ADHD at least six of the eighteen known ADHD symptoms must be met (American Psychiatric Association, 1994). Some of the symptoms refer directly to school behavior, such as "often leaves seat in classroom or in other situations in which remaining seated is expected". Parents and physicians often rely on schools and teachers since most of ADHD symptoms are exhibited almost exclusively in classroom settings. Previous studies show that schools and school teachers play an important role in identifying potential ADHD students and motivating parents to seek treatment. Sax and Kautz (2003) report that in 52.4% of the cases, ADHD diagnosis was first suggested by a child's teacher or other school personnel. In addition, DSM-IV criteria used by physicians making ADHD diagnosis relies on both parent and teacher ratings (Langberg et al., 2008). Not surprisingly, Chan et al. (2005) find that 83% of surveyed primary care physicians used teacher or school reports and ratings while evaluating a child for ADHD. Schools may also pressure parents to pursue treatment. Arcia et al. (2004) find that school reports of behavioral difficulties and direct school referrals for assessment and treatment of disruptive behavioral problems were major determinants of entry into services. In addition, pressures from schools such as insistent (e.g. daily) calls from teachers or administrative staff and suggestions to consult a pediatrician or psychiatrist were final motivators for treatment-seeking in 17.7% of interviewed mothers. Similarly, Schneider and Eisenberg (2006) found that teacher's characteristics significantly affect the probability of an ADHD diagnosis.

3. Data

Our data come from two primary sources: (1) Data on aggregate consumption of all ADHD related drugs by geographic areas provided by NDCHealth and covers the periods 1999 through 2003. (2) The National Survey of Children's Health, a nationally representative sample of individual data on children aged 0-17 collected in 2003. This data set indicates if these children are diagnosed with ADHD and whether they are currently on any medication for treating ADHD symptoms. We supplement these two primary data sets with data on timing of specific provisions of school accountability laws and accountability indexes between 1999 and 2003 at the state level. We code the accountability laws using three alternative sources: (a) Quality Counts data published on the www.edweek.org website, (b) an accountability index from Carnoy and Loeb (2002) and, (c) an accountability indicator from Dee and Jacob (2009). We use the quality counts series because it is complete and most accessible. Nonetheless, quality counts coding is non-standard and does not always match the way accountability laws have been coded in the education literature. Hence, we use two additional indexes from the literature. The differences in coding across the three sources are described below and in our analysis we provide results based on all three sources.

3.1. **Data on State Laws and Accountability Indexes.** Since the passage of the No Child Left Behind Act all states have ratcheted up their accountability efforts, though certain areas have seen more movement than others. By 2004, all states provided school report cards (although only 40 states did so in 1999), which commonly include student test scores broken down by race, family income, limited English proficiency, and disability. On the basis of these report cards, states issue public ratings of schools that identify low-performing and, in some states the highest-performing schools. Assistance to low-performing schools usually comes in the form of expert advice and a school-improvement plan. Based on ratings under the state accountability system, states may choose to offer monetary rewards to successful schools. Some states also have time limits on how long a school can be identified as low-performing before the state must take action in a form of sanctions. States have the legislative authority to withhold funds, close, take over, or "reconstitute" a failing school as a charter school. Such actions mean that the school is closed and then reopened under new management and with substantially different staff.

Previous research uses different definitions of school accountability to examine the impact of stricter accountability standards before the passage of NCLB. These different definitions fall into two broad categories. The first category separates "consequential accountability" (CA) states from "report cards" states. These studies use a categorical variable to indicate whether a state adopted some consequences based on report cards (Hanushek and Raymond, 2005, Dee and Jacob, 2009). The second group of papers develops an index to estimate the strength of consequences rather than the mere presence of consequences and assigns states to mild, moderate or strong consequences categories (Finn and Kanstoroom, 2001, Carnoy and Loeb, 2002, Lee and Wong, 2004). In this study we use three alternative indexes to measure school accountability pressures to capture both the presence and the strength of state laws.

Quality Counts Index (QCI). The Education Week Research Center's annual state policy survey tracks several individual state laws and provides them in the Quality Counts Series. For our analysis, we use five specific laws. These are, (1) Report Card: State has a report card for each of its schools? (2) Ratings: State rates schools or identifies low-performing schools? (3) Rewards: State provides monetary rewards to successful schools? (4) Assistance: State assists schools it names low-performing? (5) Sanctions: State authorized to close/takeover/reconstitute failing schools? Note that very few states have actually ever sanctioned schools. While all states now provide technical help and impose sanctions for Title I schools that fail to make adequate yearly progress, as required under the NCLB, however, the Quality Counts series credits a state *only if* their technical assistance, sanctions, and rewards apply to all public schools in the state, not just Title I schools. In 2003, thirty-six states made technical assistance available to all low-performing schools. Twenty-seven applied sanctions to such schools. Sixteen states offered rewards to high-performing or improved schools.

Based on the Quality Counts series, we construct a simple accountability index, referred to as QCI in the results, that measures the strength of accountability pressures by simply adding dummy variables for rewards, assistance and sanctions as coded in the Quality Counts series. We choose these three rather than all five because (i) the other two (report cards and ratings) are present in almost all states by 2003 and (ii) because report cards and ratings can be considered as state wide testing requirements and prior literature has made the distinction between report cards states and states that attach consequences based on report cards and ratings (Hanushek and Raymond, 2005, Dee and Jacob, 2009).

Carnoy and Leob (2002) Index (CLI). Our second index (henceforth CLI) is based on the methodology developed by Carnoy and Loeb (2002). Appendix A presents how the index was constructed. Unlike a simple sum above that counts the number of accountability laws adopted by states, Carnoy and Loeb (2002) assign a higher index to states that impose stronger sanctions (such as school closures, loss of students and school reconstitution). In addition, Carnoy and Loeb (2002) take into account student accountability as well as school accountability. Many states put pressure not just on schools but on poorly performing students who need to achieve a certain proficiency level on exit tests to graduate. Student accountability may put additional pressure on students and parents to diagnose ADHD and seek treatment to improve academic performance since consequences for poor performance in such states fall on families as well as schools. The index varies from zero (states do not conduct state-wide testing) to five (states adopted strong sanctions for poorly performing schools and require an exit test for students to graduate).

Dee and Jacob (2009) Indicator (DJI). Finally, for our third index (henceforth DJI) we use the consequential accountability (CA) approach developed by Hanushek and Raymond (2005) and used by Dee and Jacob (2009) and Wong et al. (2010). The definition of CA varies over different studies but generally includes ratings, moderate consequences (such as assistance, audit, etc.), rewards as well as strict sanctions (school closures, student transfer and reconstitution).[5] CA does not measure the strength of consequences and labels all states with mild and strong consequences as CA states. Although the CA indicator was used to asses the impact of accountability laws before NCLB it has limited use after 2002. NCLB required all states to attach a variety of consequences

[5]For instance, Wong et al. (2010) combine the CA definition used by Dee and Jacob (2009) with state proficiency rates on NAEP (The National Assessment of Education Progress) since consequences would apply to more schools if a state adopted higher proficiency standards.

to performance and thus all states effectively became CA states as soon as they phased in NCLB (Hanushek and Raymond, 2005). Even though all states adopted school ratings based on report cards, such consequences are generally considered weak (Dee and Jacob, 2009). Nonetheless, we use this indicator as given in Dee and Jacob (2009). Specifically, Dee and Jacob (2009) provide for 30 states the year a state became a CA state and we use that information to create a 1/0 indicator by state and year. Further, with the passage of NCLB, all states are coded as CA states in 2003. Because of lack of variation in this index, it is not used in the individual level analysis which uses cross-section variation in individuals diagnosed with ADHD or on ADHD medication in 2003.

3.2. **Data on Aggregate Consumption.** The consumption data is derived from NDCHealth's proprietary Source Territory Manager ® data files for the calendar years listed above which report the total number of pills sold at retail centers by strength (in milligrams) for ADHD drugs at the 5-digit zip code level within the entire continental U.S. Thus, for instance, the data set includes observations on drugs that were already on the market at the beginning of the study period (e.g. Ritalin and Ritalin-SR) as well as those that were introduced during the study period (e.g. Ritalin-LA which was introduced in 2002). The drugs that we included for our analysis were all brand names and their generic equivalents that contained either Methylphenidate HCL (MPH), Mixed Amphetamine Salts (MAS), Dextroamphetamine (DEX) and Atomoxetine (ATM).[6] For each drug, we first aggregated the data by strength within each zip code and then aggregated it up to the state level. Note that Atomoxetine is a non-stimulant molecule and was introduced in December 2002 by Eli Lilly. It attained a significant market share in 2003 (about 15%), perhaps precisely because it is the only non-stimulant ADHD drug on the market (Bokhari and Fournier, 2010). In order to insure that our results are not being driven by the introduction of this drug, we repeated all analysis by excluding this drug as well.

In addition to the drugs listed above, NDCHealth data also includes sales of two other molecules: modafinil (Provigil) and pemoline (Cylert, generics). While pemoline and modafinil are both stimulants as well, but because of their severe sides effects, neither is considered a first line drug for ADHD and are often used for treating narcolepsy. For instance, Cylert (pemoline) comes with the requirement that a prescribing physician obtain written consent from a patient prior to prescribing this drug, and specifically mentions on the label that it should not be considered as a first line therapy for ADHD. Similarly, while Modafinil is approved by the FDA for narcolepsy and a few other uses, it is not an FDA approved drug for ADHD. However, some physicians do prescribe it for ADHD as well. Thus, because of the unique nature of each of these drugs, we excluded these two molecules from the main analysis.

[6]MPH includes Ritalin, Ritalin SR, Ritalin LA, Methylin, Methylin ER, Metadate ER, Metadate CD, Concerta and generics; MAS includes Adderall, Adderall XR and generics; DEX includes Dexedrine, Dexedrine SR, Dextrostat, and generics; ATM includes Strattera.

We validated NDCHealth's data by comparing the reported quantity sold (in gms) in 1999 with the Drug Enforcement Agency's (DEA) ARCOS data for 1999, which records the total amount shipped to each area. We found the correlation between these measures from two different data sources to be very high (ranging from 0.65 to 0.9 depending on whether we compared the overall rates or individually for all counties within a state or census division) and thus provides us with reasonable confidence in our data. However, one shortcoming of this data set is that it does not tell us who is consuming these drugs. Clearly, not all of these drugs are consumed by children. An auxiliary data set obtained from the Department of Justice of California (which maintains California's data on prescriptions of Schedule II drugs due to the State's Monitoring Law) indicates that in 2001 about 67% of all psychostimulants (Methylphenidate HCL, Mixed Amphetamine Salts and Dextroamphetamines) were consumed by individuals aged 20 or less. While this figure is likely to be different across states (and somewhat over the years as well), nonetheless, as long as this percentage does not vary too much across states, we can use this aggregate data in our reduced form analysis.

3.3. **Data on Individual Diagnosis and Medication.** In addition to the aggregate data described above, we also use individual level data which identifies school aged children who have been diagnosed with ADHD, and among those diagnosed with ADHD, which students are currently on medication therapy specifically for ADHD. The National Survey of Children's Health (NSCH), collected in 2003-04, is a nationally representative individual level data on 102,353 children from 50 states and District of Columbia. Since the state in which the child resides is known, we link this data to the state accountability laws and indexes for 2003 discussed above and after controlling for individual characteristics, assess the impact of these laws on the diagnosis of ADHD and on medication therapy for a child.

4. School Accountability, ADHD Diagnosis and Medication

4.1. **Aggregate Consumption.** Between 1999 and 2003, the average consumption rate across states grew from approximately 14,321 gms/100K Children (age ≤ 20) to about 25,512 gms/100K Children. See Table 1 for a summary of mean consumption rate by year and the variation in school accountability state laws.[7] In terms of the total consumption, our data shows that sales of these ADHD drugs increased 1.8 fold between 1999 and 2003. However, other studies have documented similar or even higher increases in earlier periods. In a review of methylphenidate usage, Safer et al. (1996) report a 2.5-fold increase between 1990 and 1995 while Olfson et al. (2002) report that

[7]In the regression analysis to follow, we use log of total quantity not rates. However, for the purpose of descriptive statistics *only*, we compute and tabulate consumption rate among children by state-year as equal to $.7\times$ (total consumption in state-year) $\div$ (number of persons age ≤ 20). The 70% figure is based on the auxiliary data set by DOJ California that suggests that about 70% of all psychostimulant drugs are consumed by individuals age 20 or less.

between 1987 and 1996 stimulant use increased from 0.6 to 2.4 per 100 children. Similarly, Safer and Krager (1988) report that stimulant treatment for ADD youths doubled every 4 to 7 years between 1971 and l987.

Over these years the number of states that adopted these laws grew substantially. Per the Quality Counts (QC) series, in 1999 the number of states that had a law about issuing report cards was 38, but by 2003 almost all states had adopted such a law. Further, the number of states adopting these laws did not increase monotonically over the years. For instance, states that reward successful schools grew from 13 in 1999 to 20 in 2000 and then slowly declined back to 16 by 2003. There are also some differences in how Quality Counts codes specific laws versus how they are coded in Dee and Jacob (2009). For instance, in 1999 the number of states that had a law about imposing sanctions on poor performing schools was 18 and by 2003 the number increased to 26. By contrast, the Dee Jacob index (DJI) is one for 30 states in 1999 and is equal to one for all states in 2003.[8]

TABLE 1. Summary of State Laws and Consumption Rates[a]

	State Laws Per Quality Counts (QC)[b]					Accountability Indexes[c]			Consumption Rates[d]	
Year	R-Cards	Ratings	Rewards	Assistance	Sanctions	QCI	CLI	DJI	Mean	Stdev
1999	38	21	13	20	18	1.063	2.375	(30) .6250	14,321	3,474
2000	43	27	20	27	14	1.271	2.458	(33) .6875	15,492	3,741
2001	40	28	18	28	20	1.375	2.698	(33) .6875	16,908	4,777
2002	45	29	17	27	22	1.375	2.813	(34) .7073	19,612	5,568
2003	47	48	16	34	26	1.583	3.208	(48) 1.000	25,512	7,085
2003									(21,074)[e]	(6,017)[e]

[a]Data from 48 states (D.C., Alaska and Hawaii not included).
[b] Count of states where law is present.
[c] Mean value of the Accountability Indexes: QCI is the Quality Counts Index and is is the simple sum of dummy variables for Assistance, Sanctions and Rewards (range 0-3), CLI is the Carnoy and Loeb (2009) Index (range 0-5) and DJI is a 1/0 dummy variable, from Dee and Jacob (2009). The number in parenthesis is the number of states with Dee-Jacob accountability index equal to 1. The correlation between QCI and CLI is 0.88, between QCI and DJI is .65 and between CLI and DJI is .59.
[d]Consumption Rate computed as 70% of total consumption (in gms) divided by 100K Population Age $\leq$ 20
[e]The second set of consumption rate for 2003 excludes Atomoxetine (Strattera), the only FDA approved non-stimulant for ADHD introduced in 2003 and a blockbuster drug.

There also appears to be some correlation between state level consumption rates and accountability laws. For instance, the five-year average consumption rate in states that do not have a law about assisting poor performing schools is 17,114 gms/100K Children (or 16,551 without Strattera), while the average is 19,330 gms/100K Children (or 18,193 without Strattera) in states where such a law exists (see Table 2). Similarly, the five-year average for states without a law about imposing sanctions on failing schools is 18,001 gms/100K Children while in states when such a law is present the

[8]As noted earlier, the difference arises because QC credits a state if the law applies to all schools and not just Title 1 schools whereas the Dee Jacob index takes a value of 1 in 2003 for all states as required under NCLB.

average consumption rate is 18,885 gms/100K Children. Further, the five-year average consumption rate is always higher in states where an accountability law is present except in the case of a law about rewarding successful schools. However, the difference in five-year averages is somewhat misleading in terms of correlations since the number of states where a specific law is present is not constant over these years (as can be seen in Table 1). In fact, if we look at the year-by-year difference in averages between states that have a specific accountability law versus those that do not (see Table 2), no clear pattern emerges that would indicate presence of any strong correlation.

TABLE 2. Mean Consumption Rate[a,b] by State Law Status

	States that HAVE NOT adopted the law					States that HAVE adopted the law				
Year	R-Cards	Ratings	Rewards	Assist	Sanction	R-Cards	Ratings	Rewards	Assist	Sanction
1999	13,860	14,730	14,859	14,436	13,946	14,442	13,795	12,874	14,159	14,945
2000	13,925	15,233	15,808	15,233	15,572	15,674	15,694	15,050	15,694	15,300
2001	17,265	16,417	17,062	16,417	16,488	16,837	17,259	16,652	17,259	17,497
2002	19,569	19,012	19,492	18,275	19,798	19,615	20,005	19,831	20,652	19,392
2003	32,612	.	25,505	24,541	27,084	25,361	25,512	25,528	25,912	24,182
2003[c]	(29,009)	.	(21,231)	(20,361)	(22,426)	(20,905)	(21,074)	(20,760)	(21,368)	(19,930)
All Years	16,210	16,174	18,557	17,114	18,001	18,643	19,617	18,020	19,330	18,885
All Years[c]	(16,077)	(16,174)	(17,681)	(16,551)	(17,269)	(17,660)	(18,225)	(17,112)	(18,193)	(17,780)

[a]Data from 48 states (D.C., Alaska and Hawaii not included).
[b]Consumption Rate computed as 70% of total consumption (in gms) divided by 100K Population Age ≤ 20
[c]The second set of consumption rate excludes Atomoxetine (Strattera), the only FDA approved non-stimulant for ADHD introduced in 2003 and a blockbuster drug.

Looking at these aggregate statistics one could even make the case that it is precisely the states with a greater propensity to consume psychostimulants that pass school accountability laws. While this is certainly possible, we do not think that the passage of school accountability laws is endogenous to the consumption of psychostimulants. Thus, we employed reduced form regression analysis where we exploit the *timing* of the state laws to identify whether the adoption of accountability state laws leads to greater consumption of psychostimulants. We use variation between states and over time as accountability standards change to identify the impact of each accountability tool on psychostimulant use. Specifically, using state and year fixed effects, we regress the log of total quantity (not the rate as computed above) on a set of dummy variables for accountability laws or on the accountability index and other state level control variables and estimate regressions of the form

$$lnQ_{it} = \beta_0 + \alpha_j L_{jit} + \beta X_{it} + \sum_{i}^{48} s_i S_i + \sum_{t}^{5} \tau_t T_t + u_{it}. \tag{1}$$

In the equation above, L_{jit} is either a dummy indicator equal to one if the jth accountability law is in effect in state i in year t, or is equal to the value of the accountability index in the state-year,

S_i and T_t are state and year fixed effects, X_{it} is a vector of other state-year covariates, and lnQ_{it} is the log of total quantity consumed in state i and year t. Note that we are implicity assuming that after including the state and year fixed effects in the equation above, L_{jit} and u_{it} are not correlated.

Thus, using aggregate state level data on total consumption of psychostimulants for the years 1999-2003, we estimated Equation 1 separately for each of the individual state laws and the accountability indexes. The regression coefficients on the law variables or the indexes are given in columns marked (1) through (10) of Table 3 (the coefficients on state dummies are not shown). To account for arbitrary correlation of the error terms within a state over time, standard errors are clustered by state.

Columns (1) through (5) show the impact of individual laws (as coded by Quality Counts) on the consumption of ADHD drugs. Of these, report cards do not appear to have any impact on consumption rates while that of ratings is weak (p-value on rating coefficient is 0.09). While this could be because these two laws truly had no impact on consumption rates, it is more likely that our data is not adequate for identifying any effect of these laws. Specifically, we have a short panel (5 years, 48 states and D.C. not included) and there is not enough variation over time and states in these two laws: In 1999, 38 states had already adopted a law about issuing report cards and the number had grown to only 40 by 2001. With the passage of the No Child Left Behind Act, all remaining states more or less simultaneously adopted the report cards and ratings laws (by 2003, virtually all states had adopted these two laws, see Table 1) such that the variation in consumption rates across states and over time can not be separated from the adoption of these laws. By comparison, there is significant variation both over time and cross-sectionally in the other three laws (rewards, assistance and sanctions). Thus, if these laws had any impact on the consumption rates, we should be able to pick up their effect with our identifying strategy. Of these, only assistance and rewards have significant coefficients. Consumption increased by 4.2% for states with assistance, and by 3.3% for states with rewards. While the coefficient on sanctions is also large, it is not statistically significant (and unlike ratings, there is considerable variation in this variable over states and years).

Next, to access the cumulative impact of these three laws on consumption rates, in columns (6) and (7) we included, first just two laws and then all three simultaneously. The coefficients on all three decrease in magnitude (compare these magnitudes to those in columns (3) to (5)) and the coefficient on sanctions is not significant as before. Thus, even conditional on the presence of the other two laws, assistance and rewards have a statistically significant effect on the consumption rates but this does not appear to be so for sanctions.

Table 3. Regression of Log Quantity on State Accountability Laws

		Dependent Variable: Ln Qnty ADHD drugs										Dependent Variable: Ln Qnty non-ADHD drugs	
(N = 240)	Mean (Std)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Report Cards Law	0.888	0.013											
	(0.317)	(0.019)											
Ratings Law	0.638		0.027^{c}										
	(0.482)		(0.016)										
Assistance Law	0.567			0.042^{a}			0.037^{b}	0.031^{b}					
	(0.497)			(0.014)			(0.014)	(0.015)					
Rewards Law	0.350				0.033^{b}		0.026^{c}	0.026^{c}					
	(0.478)				(0.015)		(0.015)	(0.015)					
Sanctions Law	0.417					0.026		0.014					
	(0.494)					(0.017)		(0.017)					
QCI: Quality Counts Index	1.333								0.023^{a}			-0.012	
(Sum of three laws)	(1.134)								(0.0074)			(0.020)	
CLI: Carnoy-Loeb Index	2.710									0.0099^{b}			-0.0053
	(1.170)									(0.0044)			(0.015)
DJI: Dee-Jacob Index	0.742										0.007		
	(.439)										(.020)		
Year 2000	0.200	0.14^{a}	0.14^{a}	0.13^{a}	0.13^{a}	0.14^{a}	0.13^{a}	0.13^{a}	0.13^{a}	0.14^{a}	0.14^{a}	0.51^{a}	0.51^{a}
	(0.401)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.044)	(0.044)
Year 2001	0.200	0.23^{a}	0.23^{a}	0.22^{a}	0.22^{a}	0.23^{a}	0.22^{a}	0.22^{a}	0.22^{a}	0.22^{a}	0.23^{a}	0.76^{a}	0.76^{a}
	(0.401)	(0.020)	(0.020)	(0.021)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)	(0.020)	(0.073)	(0.073)
Year 2002	0.200	0.39^{a}	0.39^{a}	0.38^{a}	0.39^{a}	0.39^{a}	0.38^{a}	0.38^{a}	0.38^{a}	0.38^{a}	0.39^{a}	1.14^{a}	1.14^{a}
	(0.401)	(0.026)	(0.026)	(0.027)	(0.025)	(0.026)	(0.027)	(0.026)	(0.026)	(0.026)	(0.026)	(0.090)	(0.090)
Year 2003	0.200	0.67^{a}	0.66^{a}	0.65^{a}	0.66^{a}	0.67^{a}	0.65^{a}	0.65^{a}	0.65^{a}	0.66^{a}	0.67^{a}	1.39^{a}	1.39^{a}
	(0.401)	(0.032)	(0.034)	(0.034)	(0.032)	(0.032)	(0.034)	(0.034)	(0.033)	(0.032)	(0.033)	(0.11)	(0.11)
Ln Population	15.122	-0.88^{b}	-0.88^{b}	-0.88^{b}	-0.79^{b}	-0.91^{a}	-0.82^{b}	-0.84^{b}	-0.86^{b}	-0.88^{b}	-0.87^{b}	0.76	0.76
	(0.992)	(0.34)	(0.34)	(0.34)	(0.34)	(0.33)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(1.10)	(1.12)
% Age 5-19	21.524	0.029	0.031	0.022	0.028	0.025	0.024	0.023	0.023	0.021	0.029	0.013	0.015
	(1.293)	(0.020)	(0.020)	(0.020)	(0.021)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.080)	(0.081)

Notes: All regressions include state dummies. The sample consists of 48 States (Hawaii, Alaska and D.C. are excluded). Standard errors (clustered by state) are in parenthesis and a, b, c are significance levels at 1,5 and 10% respectively.

We further verified the cumulative effect of these three laws by constructing a simple accountability index as the sum of the dummy variables (where the sum ranges from zero to three) and using that in the regression analysis. The results are given in column (8) and show that a one point increase in the index is associated with a 2.3% increase in the consumption.[9] Thus, having two or three of these laws on the books is associated with a significant increase in psychostimulant consumption relative to states that enforce none or just one of these three laws. We also verified the impact of the strength of these laws by using the alternative accountability index CLI (the Carnoy-Loeb Index). The results are given in column (9) and show that a one point increase in the CLI index is associated with a 0.99% increase in the consumption. Further, the two latter coefficients are comparable in magnitudes: multiplying the estimated coefficients with their respective mean values gives elasticity of consumption (at the sample mean) with respect to the first index as .031 and .027 with respect to the CLI index.

Finally, when we re-estimated the model using the DJI index (the consequential accountability index of Dee and Jacob (2009)), the coefficient on it was .007 but was not statistically significant (p-value .72). Our estimation strategy uses both state and year dummies along with an accountability index and the lack of significance on DJI is due to little variation in the timing of this index over states: DJI is a dummy variable that starts with a value of 1 for 30 states in 1999, three additional states are coded as one in 2000 and then all remaining states simultaneously switch to a value of 1 in 2003 (see Table 1). We further verified that the lack of significance is due to little variation in the timing of the index by replacing the year dummies with linear and quadratic time trends (but retaining the state dummies). With just the linear time trend the coefficient on DJI Index was .079 with a p-value less than .001 and if we also included a quadratic time term, the coefficient on DJI decreased to .031 with a p-value of .093 (a polynomial of degree four in time is identical to year dummies and almost all the variation in the index is absorbed by the year dummies).

Taken in combination, these results suggest the following. The Dee and Jacob index – which captures only the presence/absence of any accountability pressures – cannot account for the rise in consumption of ADHD drugs over time but does explain some variation across states. However, the other two indexes, which are sensitive to the strength of these accountability pressures, explain both the variation over time and across states in the aggregate consumption of ADHD drugs.

A Falsification Test. We repeated the analysis using data on pemoline and modafinil consumption. As mentioned earlier, these two molecules are also psychostimulants but they are not considered first

[9] We also experimented with non-linear effects in the accountability index. For instance, we constructed two alternative dummy variables indicating if there were two or three out of three laws in effect (the comparison group was zero or one law in effect) and found that presence of any two such laws is associated with 2.3% increase in the consumption while having all three laws in effect is associated with a 6.7% increase in the consumption rates. Similarly, we also estimated the model with zero and one law separated out into two separate categories. The results were similar though somewhat weaker since the panel is not large enough to allow such refined division of observations into that many cells.

line drugs for ADHD. While the total consumption of the main ADHD drugs (MPH,MAS,DEX,ATM) grew by 80.5% over 1999-2003 period, the total consumption of pemoline and modafinil grew by 312% over the same years. However, as columns (11) and (12) show, the coefficient on the accountability index (QCI or CLI) is not significant. Similar result holds for individual laws and the third index (not shown). These additional regressions indicate that our law dummies or the accountability index in the main analysis are not just picking up the effect of other state level variables that may be correlated with both the laws and the consumption of ADHD drugs.

A Short Panel Limitation. Both the descriptive statistics as well as the coefficients on the year dummies in Table 3 show that the rate of psychostimulant consumption has increased remarkably over the short panel of the data. Additionally, these are the same years when accountability laws were ratcheted up. Ideally, one would like to see evidence from longer-term state-specific trends prior to the ramping up of accountability laws.[10] However, longer time series on consumption of ADHD drugs is not available. Absent the longer panel, as a second best option we used state-level data from the Common Core of Data (aggregated up from the school district level) from the pre-1999 period to come up with a pre-1999 trend in state-level fraction of students with Individualized Education Plans (IEP), and then controlled for predicted values from this trend in the post-1999 period.[11] Thus, we first computed the fraction of all school children with IEP for each state for the years 1993-1998. We then used state specific trends to predict the values of IEP for each state for the years 1999-2003 and repeated the analysis in Columns (1) through (10) of Table 3 but included this predicted value of IEP as an additional control. The results were very similar to the ones reported above. For instance, after controlling for predicted value of IEP, the coefficient on QCI index was .022 with the p-value of .01 (compare this to the coefficient of .023 in column (8)) and the coefficient on CLI index (Carnoy and Loeb index) was .0091 with a p-value of .061 (compared to .0099 reported earlier in column (9)).

Additional Robustness Checks. In the preceding regressions, our identifying strategy was to make use of the variation in the timing of adoption of laws by states. Further, we have relied on state and year fixed effects (as well as the size of the relevant population) to capture the effect of all remaining variables that are either common to all states and changing over time or common to all

[10]We are grateful to an anonymous referee for pointing out this shortcoming as well as for suggesting the second best option.

[11]Individuals with Disabilities Education Act (IDEA) explicitly incorporates ADHD within the definition of "Other Health Impaired" which makes them eligible for special education services. According to the U.S. Department of Education guidelines, each public school child must have an Individualized Education Program (IEP) written within 30 calendar days after a child is determined eligible for special education (U.S. Department of Education, 2000). Thus, to qualify for the IEP, ADHD students must show that their disability adversely affects their educational performance and that they need special education. Since ADHD and learning disability are often comorbid, previous literature shows that between 26.6% and 55.4% of ADHD students receive IEP (Smith and Adams, 2006) and such plans have been shown to be important moderators of medication therapy (Scheffler et al., 2009). In addition, previous literature shows that schools may respond to higher accountability pressures of showing academic progress of students with disabilities through involvement of the IEP team (Browder and Cooper-Duffy, 2003). Therefore, we believe that a rise in IEPs is a close proxy for the rise of ADHD in school-aged children.

time periods but are state specific. However, state specific factors that are changing over time are not controlled for in the regressions and are still in the error term. If these variables are correlated with the accountability laws, then the estimated coefficients on the state accountability laws are biased (and inconsistent).[12] To check if this is true, i.e., if the unobserved state specific factors that are changing over time are correlated with the accountability laws, we estimated additional fixed effects regressions but this time included a series of covariates specific to states and changing over time, i.e., for each of the regressions above, we included a vector of variables X_{it}. Specifically, we included log of total population, percent of African-American population, percent of other minorities, percent of population aged 5 to 19, log of children enrolled in special education programs, log of child population participating in school lunch programs, student teacher ratios, unemployment rates, per capita income, percent uninsured, log of S-CHIP compensation and log of total medicaid population in the state.

With the exception of the regression on the ratings law (column 2 earlier, which had a p-value of .09), adding in these covariates did not change the coefficients on the state law dummies appreciably (in magnitude) in any of the other regressions, nor did the level of significance ever change. The results from these additional regressions are given in the appendix (see Table B-1). While this is far from strong statistical evidence in support of no omitted variables *bias*, it does increase confidence in our fixed effects approach to the extent that if these other unobserved omitted variables are similar to the ones that were included, they are not necessarily correlated with the state law dummies. In fact, the only time we see an appreciable change in the magnitudes of the law dummies is if we do not include any state dummies. This gives us some confidence that any state-time covariates that are omitted from Equation 1 and which may be correlated with the timing of the law are absorbed in the state and year dummies. Thus, including state and year dummies reduces or eliminates omitted variables bias.

Finally, we repeated the entire analysis (with and without the covariates), by removing Atamoxetine from the aggregate consumption variable. Results without the inclusion of this drug were very similar to the ones reported above have been omitted.

4.2. **Probability of Diagnosis and Medication Therapy.** A shortcoming of the above analysis is that we actually do not know what fraction of these psychostimulant sales was for school aged children. As long as either the fraction of total sales to children is the same across states, or if it is different, then it is not systematically correlated with the accountability laws, the above analysis

[12]Two common sources of such endogeneity are omitted variables and simultaneity (or reverse causality). We do not think that simultaneity is a realistic concern since that would require that state legislators are passing accountability laws at specific points in time in response to consumption of ADHD drugs. However, omitted variables bias is still possible. The standard method to correct for such a bias would be through the use of instrumental variables (or to check for it via the usual Hausman test). We do not have any valid instruments (i.e., uncorrelated with consumption rates) that also correctly predict the timing of accountability laws by state (i.e. are also relevant).

is still valid (since then it is absorbed by the state dummy variables). Thus, the preceding analysis assumes that the differences in the fraction of sales to children across states is not correlated with the accountability laws. However, no such assumption in needed in the individual level analysis where we know if a child is diagnosed with ADHD and if they have been prescribed medication specifically for ADHD.

ADHD is not a discrete medical condition, though the diagnosis is discrete, and hence we investigate the impact of accountability laws on ADHD diagnosis and medication therapy via the use of a latent variables model. Since medication can only be prescribed if a child is first diagnosed with ADHD, we estimate a sequential probits model of the form $D_i = 1$ if $D_i^* > 0$ and $(M_i|D_i = 1) = 1$ if $M_i^* > 0$ where

$$D_i^* = \beta_1 L_i + \sum_{j=2}^{k} \beta_j X_{ji} - \epsilon_{1i} \tag{2a}$$

$$M_i^* = \alpha_1 L_i + \alpha_2 \tilde{D}_i + \alpha_3 L_i \tilde{D}_i + \sum_{j=4}^{k} \alpha_j X_{ji} - \epsilon_{2i} \tag{2b}$$

and $\tilde{D}_i = Pr(D_i^* > 0)$.

In the first equation above, D_i^* is the *perceived* severity of ADHD of child i, as perceived by a physician, and is a function of individual characteristics of the form $\sum_{j=2}^{k} \beta_j X_{ji}$ of the ith child. While the diagnosis of ADHD is given solely by the physician based on her/his own observation of the child, in turn the physician also relies on symptoms and behavior as reported by parents and teachers. Since in the presence of accountability laws, the latter group *may* have an incentive to have a child diagnosed with ADHD, they can influence the physician's perception of the severity of ADHD (via the reports of classroom behavior and symptoms) and hence we also specify D_i^* as a function of L_i, the accountability index value in the state of student i. If $\beta_1 = 0$, then accountability laws do not influence the perception of severity of ADHD and consequently the decision to diagnose a child as ADHD.

In the second equation, M_i^* is the *expected* efficacy of medicating a child and is a function of his/her characteristics and of the probability that the child is diagnosed with ADHD. The higher the probability of diagnosis, the greater is the underlying severity of ADHD, and hence greater the efficacy of medicating a child. Thus, we expect $\alpha_2 > 0$. Once again, the final decision to start medication therapy is reached jointly only by the physician and parents, but the decision process is not without input from the teachers (at least indirectly) and hence the expected efficacy of the medication therapy may also be a function of school accountability laws. We test for this by letting M_i^* be a function of L_i and of the interaction of L_i with $\tilde{D}_i$.

Since the model specified above is a sequential probit (a person can be medicated only after they have been diagnosed) it is internally consistent and the parameters are identifiable.[13] Thus, our estimation strategy relies on two parts. In the first part, we use the entire sample and estimate the first equation via a probit and where the coefficient on the indicator variable tells us if the accountability law has an impact on the probability of receiving an ADHD diagnosis. Next, we use the coefficients from this probit to compute $\widehat{\widehat{D}}_i = \Phi(X, L; \widehat{\beta})$ and use these in the probit estimation of the second equation which is estimated on the sub-sample of children diagnosed with ADHD. The coefficients α_1 and α_3 identify the impact of the accountability laws on the probability of receiving medication therapy, conditional on ADHD diagnosis.

Sample. The medical and epidemiological literature lists several risk factors associated with the diagnosis of ADHD and drug therapy, including (but not limited to) age, gender, race, ethnicity, income, insurance, parent's education and family structure. In our estimation of the probits specified in Equation 2, we include all these variables in the vector X_{ji}. Additionally, we also include some of the state level variables used in the aggregate analysis (percentage of population aged 5 to 19, log of children enrolled in special education programs, log of child population participating in school lunch programs, student teacher ratios, log of S-CHIP funds, log of medicaid population and log of S-CHIP population). For our analysis, we restricted the sample to school-aged children (ages 5 to 17) who are enrolled in public schools (those in private schools are not subject to the accountability laws). The final sample, and for which the information on the covariates in not missing, consists of 49,527 children. Among these sample children, 4,715 (9.52%) were diagnosed with ADHD while the remaining 44,812 (90.48%) were not. Further, of the 4,715 children with ADHD, medication therapy status is known for all but 14 children. Thus, of the 4,701 children, 2,824 (60.1%) were on ADHD medication therapy. Data for selected covariates is summarized in Table 4.

Descriptive Statistics. The NSCH data set is based on a complex survey design and provides variables that identify stratas (states) and the post-stratification probability of including a child in the sample (weights). Thus, from here onwards, we only report results that weight the observations appropriately and where standard errors are always clustered by states. While only 9.52% of all children in our sample are diagnosed with ADHD, 85.5% of them reside in states where the assistance law was in effect. By comparison, 82.9% of those not diagnosed with ADHD reside in similar states. Further, 37.4% and 73.2% of children diagnosed with ADHD live in states with rewards and sanction laws but among those who are not diagnosed with ADHD, 32.6% and 72.5% live in states with similar laws. In each of these cases, a slightly higher percentage of children diagnosed with ADHD live in states with these three accountability laws compared to the percentage of children not diagnosed with ADHD. Similarly, the mean values of QCI and CLI indexes are higher for children diagnosed

[13] See Maddala (1983) pp. 123 for a similar model (eqn. 5.51).

with ADHD compared to their counterparts. Of those diagnosed with ADHD, 60.07% are on drug therapy for ADHD. However, no clear pattern on medication status by state law exists for these children: a smaller fraction of children on medication therapy live in states that have assistance and sanctions laws (84.5% and 71.8% respectively) compared to the percentage of children not on medication therapy (86.8% and 75.0%), but a slightly larger fraction of children on medication therapy reside in states that have the rewards law in effect compared to the fraction of children not on medication therapy (38.1% vs. 36.5%).

Table 4. Proportion of Children in States with Accountability Laws by ADHD and Medication Status

	ADHD? (N=49,527)		Medication? (N=4,701)| ADHD=Yes	
	Yes(9.52%)	No(90.48%)	Yes(60.07%)	No(39.93%)
Assistance Law (1/0: 1 if State Law in effect)	0.855 (0.053)	0.829 (0.061)	0.845 (0.058)	0.868 (0.048)
Rewards Law (1/0: 1 if State Law in effect)	0.374 (0.092)	0.326 (0.087)	0.381 (0.093)	0.365 (0.095)
Sanctions Law (1/0: 1 if State Law in effect)	0.732 (0.073)	0.725 (0.075)	0.718 (0.077)	0.750 (0.071)
QCI: Quality Counts Index (Sum of the three laws)	1.962 (0.140)	1.880 (0.137)	1.945 (0.142)	1.984 (0.142)
CLI: Carnoy-Loeb Index	3.942 (.209)	3.848 (0.215)	3.935 (0.227)	3.947 (.190)
Age	11.911 (0.072)	11.542 (0.033)	11.362 (0.095)	12.665 (0.096)
Gender (1/0: 1 if Male)	0.729 (0.009)	0.492 (0.004)	0.733 (0.012)	0.723 (0.016)
Gender (1/0: 1 if Female)	0.271 (.009)	0.508 (.004)	0.267 (0.012)	0.277 (0.016)
Race (1/0: 1 if White)	0.790 (0.013)	0.747 (0.150)	0.812 (0.122)	0.758 (0.019)
Race (1/0: 1 if African American)	0.143 (0.018)	0.166 (0.018)	0.125 (0.016)	0.168 (0.027)
Race (1/0: 1 if Other)	0.067 (0.010)	0.087 (0.017)	0.063 (0.008)	0.074 (0.015)
Ethnicity (1/0: 1 if Hispanic)	0.053 (0.012)	0.089 (0.022)	0.042 (0.008)	0.068 (0.023)

Except for Age, he table displays the mean of the dummy variables. Clustered standard errors in parenthesis

Selected descriptive statistics for a a few of the typical 'risk factors' associated with ADHD diagnosis and medication therapy are summarized in Table 4. Observe that children diagnosed with ADHD are slightly older than their counterparts while those on medication (conditional on being diagnosed) are younger than those not on medication. Similarly, compared to females, males are significantly more likely to be diagnosed with ADHD (among those diagnosed, 72.9% are males while only 27.1% are females) as well as be on medication for ADHD (among those medicated, 73.3% are males while only 26.7% are females). However, while the conditional distributions of gender, conditional on ADHD = Yes/No are very different, the conditional distributions of gender, conditional on Medication = Yes/No (and ADHD=yes) are very similar (note that the mass of the conditional distributions on the last two columns under medication status for gender are very similar but those under diagnosis are very different). Thus the conditional distribution of gender, conditional on medication status, appears to be orthogonal to the medication and suggests that while gender is a significant risk factor for diagnosis, there is little difference in medication therapy by gender among those diagnosed. Finally, both race and ethnicity appear to be significant risk factors for the diagnosis of ADHD (the conditional distributions by ADHD status are quite different): compared to White children, African Americans and others are less likely to be diagnosed with ADHD. Similar differences in race/ethnicity conditional distributions appear by medication status, but since it is also conditioned on ADHD status equal to yes, it is not clear if the differences in conditional distributions of race/ethnicity by medication status are merely a reflection of the difference already noted in diagnosis status.

Probability of Diagnosis. While these descriptive statistics are revealing, they do not provide the full story of how these laws affect the probability of diagnosis and medication for children who have ADHD. Thus, we estimated the probability of diagnosis and subsequent medication per specification given in Equation 2. Columns (1) through (6) of Table 5 provide the estimated coefficients on accountability laws/indexes of the probability of being diagnosed with ADHD (the table displays only selected coefficients and the full set of results is given in the appendix in Table B-2).

In the first three specifications, we included the three accountability laws one at a time in the probit for ADHD. Column (1) shows that the coefficient on the assistance law is .13 and is statistically significant at the 1% level. The marginal effect of the law on the probability of being diagnosed is .0181 at the sample mean with the associated p-value less than 0.000 (the mean of the marginal was also computed and was .0182). When we replace the assistance law with the rewards law, the coefficient is slightly smaller (.076) and is significant at the 5% level. The marginal effect of the law at the sample mean is .012 with a p-value of .051 while the mean of the marginals is .011. Finally, when we use the sanctions law instead, the coefficient is .036 and is not statistically significant. The marginal effect for this third law is .0054 (at the sample mean while the mean of the marginals is .0052) with a p-value of .381. Thus, the first two laws appear to significantly increase the probability of diagnosis while the third does not have a significant impact on the probability of diagnosis.

Two things are worth noting about these three results. First, these results are consistent with earlier results on aggregate consumption of psychostimulants where, for instance, assistance had the largest coefficient and was significant at $p < .01$ followed by the coefficients on the other and sanctions was not significant (see Table 3 columns (3),(4) and (5)). Second, the signs of these coefficients are also consistent with the earlier descriptive statistics in Table 4. Even the lack of significance on the sanctions law is somewhat predictable, given the small difference in the conditional mean value of the sanctions law (.732 vs. .725) seen in Table 4.

Table 5. Probability of Diagnosis

Selected coefficients shown. Detailed results given in the appendix.

	Main Analysis: Public School Students (1-6)						Double Falsification Test (7-10)			
							Pvt. Sch. Students		Pub. Sch. Students 1000 Bootstraps on	
	N=49,527						N=6,714		N=6,714	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Assistance Law	0.13^{a}			0.15^{a}						
(1/0: 1 if State Law in effect)	(0.028)			(0.030)						
Rewards Law		0.076^{b}		0.079^{a}						
(1/0: 1 if State Law in effect)		(0.038)		(0.031)						
Sanctions Law			0.036	-0.046						
(1/0: 1 if State Law in effect)			(0.041)	(0.030)						
QCI: Quality Counts Index					0.055^{a}		-0.021		0.058^{c}	
(Sum of the three laws)					(0.017)		(0.037)		(0.040)	
CLI: Carnoy-Loeb Index						0.034^{a}		-0.017		0.036^{c}
						(0.011)		(0.024)		(0.029)
Age	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.056^{a}	0.056^{a}	0.015^{c}	0.015^{c}
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.010)	(0.010)	(0.011)	(0.011)
Gender (1/0)	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.527^{a}	0.526^{a}
(1 if Male)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.079)	(0.079)	(0.079)	(0.079)
Race (1/0)	-0.30^{a}	-0.30^{a}	-0.29^{a}	-0.31^{a}	-0.30^{a}	-0.30^{a}	-0.053	-0.052	-0.320^{a}	-0.317^{a}
(1 if African American)	(0.049)	(0.053)	(0.050)	(0.051)	(0.051)	(0.050)	(0.20)	(0.20)	(0.121)	(0.121)
Race (1/0)	-0.17^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.26^{a}	-0.26^{a}	-0.189	-0.191
(1 if Other)	(0.051)	(0.052)	(0.052)	(0.051)	(0.051)	(0.051)	(0.099)	(0.098)	(0.168)	(0.168)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: Columns (9) and (10) show the mean and standard deviation of the 1000 bootstraps (with replacement). The bootstraps adjust for the observation weights in the data set.

To access the joint impact of the three laws, we estimated a probit where we included a dummy variable for each of the three laws simultaneously. The coefficients on assistance and rewards are positive (.15 and .079 respectively) and significant while the coefficient on sanctions is -.046 with a p-value of .13. Since the three laws are positively correlated with each other, this last result is unexpected and difficult to interpret. However, further inspection of the data revealed that the result was driven solely by one state, Alabama. In 2003, Alabama had the highest ADHD diagnosis rate in the nation, 12.67% among children in public schools (compared to the national average of 9.53%) and was also the *only* state where assistance and rewards laws were in effect but sanctions were not. Thus, in our sample, the percentage of children who are diagnosed with ADHD when all three laws are in effect (which is 13 states) is 10.84% while the percentage of children diagnosed with ADHD when assistance and rewards are in effect but not sanctions (which is just one state, i.e. Alabama) is 12.67%. To verify that the negative and nearly significant coefficient on sanctions (conditional on assistance and rewards) is due to Alabama, we re-estimated the specification excluding observations from Alabama. The results for the dummy variable on rewards and assistance remained unchanged (both in magnitude and significance) while that for sanctions decreased in magnitude and had a p-value of .28.

Next, we repeated the analysis using the two accountability indexes.[14] The results are shown in columns (5) and (6). In both cases, the coefficient is positive and significant. The marginal effect at the mean is .0083 (p-value = .002) for QCI and .0052 (p-value = .002) for CLI (the mean of the marginals for the two indexes are .0078 and .0049). Again these result mirror those in columns (8) and (9) in Table 3.

Finally, observe that in all six specifications, the coefficients on age and gender (male) are positive and significant while those on race (African Americans and Others) and Ethnicity (Hispanic) are negative and significant and do not change across the specifications. These results are consistent with both the descriptive statistics as well as the (voluminous) literature on risk-factors associated with ADHD (see Bokhari et al. (2008)).

A Falsification Test. In order to check if accountability indexes are truly capturing the effect of accountability laws on the probability of being diagnosed, rather than some other unobserved factors at the state level correlated with ADHD diagnosis, we re-estimated the specifications in columns (5) and (6) on a sub-population where these laws should *not* have any effect: children enrolled in private schools, since the accountability laws only effect public schools. Results are given in columns (7) and (8). Observe that the point estimate on the two accountability indexes is much smaller and is not statistically significant in either of the two cases. If in fact the accountability indexes were capturing the effect of other unobserved state level factors correlated with ADHD diagnosis, then we should expect to see the coefficient on these indexes to be similar to their counter parts in the public school population, which they are not. Further, among the private school students, age and gender remain significant risk factors associated with diagnosis (and the coefficient on age increases) and the dummy variables on African-Americans and Hispanics are no longer significant. While these results suggest that these accountability indexes are really capturing the combined effect of the accountability laws and not of other unobserved factors (especially since the coefficients on age and gender remain significant – consistent with the literature on predictors of ADHD), the possibility remains that the lack of significance on the accountability indexes is simply a power

[14] Note that we have dropped DJI index (the Dee and Jacob index) from the analysis since, per this index, all states are coded as one in 2003 and the NSCH sample on individuals in available only for year 2003.

issue as there are only 6,714 children in our sample for the private school population compared to 49,527 observations in the public school population. Thus, to further verify if this is indeed a power issue, we constructed 1000 random samples (with replacement) of size N=6,714 *from the public school population* and re-estimated the original probit on these sub-samples. The mean and standard deviation (i.e. the bootstrapped standard error) of the coefficients is reported in columns (9) and (10). Observe that all the point estimates, including those in the accountability indexes, are now similar to those in columns (5) and (6) (and not statistically different from them) and the coefficients on the two accountability indexes are still significant at the 10% level. Results in columns (7)-(10) together provide more confidence that the indexes are in fact capturing the effects of accountability laws and not other unobserved factors correlated with the diagnosis.

Probability of Medication. To assess the impact of these laws on the probability of receiving medication therapy, we first estimated a simple probit on the full sample without regard to whether a child is ADHD or not (the sample reduces to 49,513 observations because we do not have medication information for 14 children). The results are shown in columns (1) and (2) of Table 6 (the table shows selected coefficients, full set of results is given in the appendix in Table B-3).

First, observe that the probability of medication, unlike the probability of diagnosis, *decreases* with age (as suggested by the descriptive statistics as well, see Table 4) and that females, non-white children and those of Hispanic origin are less likely to receive medication therapy. Next, the coefficient on the two accountability indexes is positive and significant (.051 and p-value = .014 and .033 and p-value = .024). These results are consistent with the earlier aggregate analysis on total consumption where again the two indexes were significant. Further, the marginal effect (for the probability of medication) with respect to the accountability indexes is much smaller than that computed earlier for probability of ADHD. The marginal effect for the probability of medication at the sample mean is .0050 (p-value = .014 and the mean marginal effect is .0056) for the first index indicating that every one point increase in the index (on a discrete scale from 0 to 3) on average increases the probability of medication by about .005. The marginal at the sample mean for the second index (CLI – the Carnoy and Loeb index) is .0032 (p-value = .023 and the mean marginal is .0036). These marginals are understandably small in magnitude since we estimated this model on the full sample, i.e., children with and without ADHD diagnosis while the medication therapy can only be prescribed if a child is first diagnosed with ADHD.

TABLE 6. Probability of Medication

Selected coefficients shown. Detailed results given in the appendix.

	Pr(MEDS=1) (1)	(2)	Pr(MEDS=1|ADHD=1) (3)	(4)
QCI: Quality Counts Index	0.051^{b}		0.096	
(Sum of the three laws)	(0.021)		(0.077)	
CLI: Carnoy-Loeb Index		0.033^{b}		0.053
		(0.014)		(0.060)
$\widehat{\tilde{D}}$: Probability of ADHD			5.69^{c}	5.48
			(3.14)	(3.49)
$\widehat{\tilde{D}}\times$ Index			-1.02^{b}	-0.53
			(0.49)	(0.35)
Age	-0.013^{a}	-0.013^{a}	-0.11^{a}	-0.11^{a}
	(0.0043)	(0.0043)	(0.011)	(0.012)
Gender	0.47^{a}	0.47^{a}	-0.26	-0.24
(1/0: 1 if Male)	(0.032)	(0.032)	(0.25)	(0.25)
Race	-0.35^{a}	-0.35^{a}	-0.087	-0.10
(1/0: 1 if African American)	(0.056)	(0.057)	(0.15)	(0.15)
Race	-0.18^{a}	-0.18^{a}	-0.0056	-0.011
(1/0: 1 if Other)	(0.063)	(0.063)	(0.14)	(0.14)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: In column (3), $\widehat{\tilde{D}}$ and $\widehat{\tilde{D}}\times$ Index is computed using the Quality Counts index. Similarly, in columns (4) these two variables are computed using the Carnoy-Loeb index.

We finally estimated the probability of receiving medication therapy, conditional on having received an ADHD diagnosis via the specification given in Equation 2. Using probit coefficients from $Pr(D_i = 1)$, (for each of the two indexes separately) we computed $\widehat{\tilde{D}}_i = \Phi(X, L; \widehat{\beta})$ and used this value in the estimation of $Pr(M_i = 1|D_i = 1)$ on the sub-sample of children diagnosed with ADHD (N= 4,701). The variable $\widehat{\tilde{D}}_i$ can be interpreted as the estimated normalized value of the latent variable, i.e., the perceived severity of ADHD. Results are given in columns (3) and (4) but are not directly comparable to those in columns (1) and (2) because the two sets of probits are for different sub-populations (unconditional on ADHD versus those who are ADHD) and also because the specifications are different. The coefficient on the accountability index is not significant in either of the two cases while that on $\widehat{\tilde{D}}_i$ is positive but significant for only the first case (in the case of the second index, CLI index, the p-value for the coefficient on $\widehat{\tilde{D}}_i$ is .11). The mean marginal effect with respect to the the first accountability index is -.011 while the marginal at the mean is -.014 with a p-value of .48 (the computations of the marginals explicitly account for the presence of the interaction term).[15] Similarly, in the case of CLI index, the mean marginal effect was -.0048 and the marginal at the mean was -.006 with a p-value of 0.69. Thus, conditional on the severity of ADHD (or the probability of being diagnosed with ADHD) the accountability laws do not effect the probability of medication therapy. In combination with the earlier results, total medication rate increases with the accountability laws because more children are diagnosed with ADHD, but once diagnosed, approximately 60% are prescribed medication therapy regardless of accountability laws.

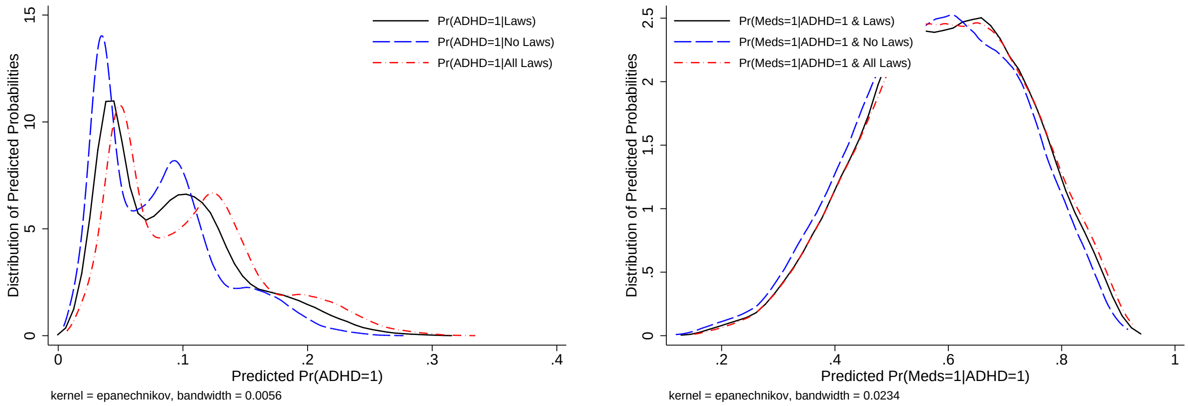
Combined Effects. All three laws were in effect in 13 states while in another 12 none of the three laws were in effect. To assess the joint impact of these laws (or lack thereof) on the probability of diagnosis as well as on medication, we computed the model-predicted probabilities.[16]. The predicted probability of diagnosis ranged from .005 to .310 with a mean value of .092 (recall that in our sample 9.52% of the children are diagnosed with ADHD, see table 4). The distribution of predicted probabilities is right-skewed as shown on the panel on the left in Figure 1.

The figure also shows the distribution of model-predicted probabilities of ADHD diagnosis for (a) if each child in the sample was in a state where there were no accountability laws, and (b) if each child in the sample was in a state where all three accountability laws were in effect. In the first case the mean probability is .079 while in the second case the mean probability is .105 (difference is .026). Since nationwide there are about 42,776,150 children in public schools, this difference in probabilities implies 1,099,415 additional children were diagnosed with ADHD due to these laws. The average value of change in probability (.026 from no laws to all laws) is the average for the entire population but is in fact larger for males (.034 compared to .017 for females) and whites (.027 compared to .022 for non-white) and lower for hispanics (.017 compared to .027 for non-hispanics). Put another way, the presence/absence of these laws does not effect all demographic groups equally. Similarly, we also computed the model predicted probability of receiving medication therapy for each child with ADHD and conditional on their severity. The predicted probability ranged from .15 to .92 with the mean value of .596 (note that 58.76% of children with ADHD receive medication therapy). Figure 1 also shows the distribution of predicted probabilities along with the predicted distributions if all laws were in effect and when none were in effect. In these hypothetical cases, there is no significant shift in the distributions when there are no accountability laws versus when all three laws are present. Thus, of the 1,099,415 additional children diagnosed with ADHD (associated

[15]Specifically, since the probability takes the form $Pr(M=1|D=1) = \Phi(\alpha_1 L_i + \alpha_2 \widehat{\widehat{D}}_i + \alpha_3 L_i \widehat{\widehat{D}}_i + \sum_{j=4} \alpha_j X_{ji})$, then the marginal was computed as $\partial Pr(M=1|D=1)/\partial L_i = \phi(\cdot) \times (\alpha_1 + \alpha_3 \widehat{\widehat{D}}_i)$. This expression was evaluated for each observation in the sample to compute the mean marginal effect. Additionally, it was also evaluated at the sample mean (to compute the marginal at the mean) and the standard error was obtained via the delta method.

[16]The computations shown here are for the first accountability index and are qualitatively similar to those from the CLI index.

FIGURE 1. Distribution of (predicted) Probabilities



	$Pr(ADHD = 1)$				$Pr(Meds = 1 \mid ADHD = 1)$			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Laws as is	.092	.054	.005	.310	.596	.141	.152	.916
No Laws	.079	.047	.004	.277	.582	.142	.119	.917
All Laws	.105	.058	.007	.334	.600	.142	.152	.930
Difference	.026	.011	.003	.058	.016	.048	-.162	.108

	Difference by Demographics			
	Mean	SD	Min	Max
Male	.034	.008	.007	.058
Female	.017	.006	.003	.038
Race (White)	.027	.011	.004	.058
Race (Black)	.022	.010	.003	.046
Race (Other)	.022	.010	.003	.052
Hispanic	.017	.009	.003	.042
Non-Hispanic	.026	.011	.004	.058

with the presence of these laws) approximately 59.6%, i.e., 655,252 are also prescribed medication therapy regardless of the accountability laws.

5. SUMMARY AND CONCLUSION

We used state level data on the consumption of psychostimulant drugs and relied on state and year variation in the provision of school accountability laws to identify the effect of these laws on consumption. Our results indicate that provisions of assistance and rewards to schools are associated with a 3.1% and 2.6% increase in consumption, and a one unit increase in an accountability index constructed using Quality Counts series (QCI, range 0-3) is associated with about 2% increase in consumption, while a unit increase in a second index (CLI, range 0-5) based on Carnoy and Loeb (2002) is associated with about 1% increase in consumption. However, sanctions and the consequential accountability index (DJI, 1/0 dummy), based on Dee and Jacob (2009) are not significantly related to aggregate consumption. While these results show that more stringent accountability laws have a significant impact on the aggregate consumption of psychostimulants, they do not reveal whether this increase is on the extensive or intensive margin, i.e., in association with the accountability laws, are more children being diagnosed with ADHD and hence the consumption rates increase, or is it that children already diagnosed with ADHD are more likely to be prescribed psychostimulant drugs? To this end, we used the NSCH data set and estimated a series of probability models. Our results indicate that these provisions increase the probability of being diagnosed with ADHD by .018 and .012 for assistance and rewards respectively (once again coefficient on sanctions is not significant) and the marginal effects of the two accountability indexes are also positive and significant. Further, conditional on ADHD diagnosis, the accountability laws do not change the probability of receiving medication therapy. Using the private school population, we also constructed a falsification test to check if other unobserved state level factors, which are correlated with accountability laws and effect the probability of ADHD diagnosis, are driving our results.

Note that in general it is assistance, rewards and the accountability indexes QCI and CLI (which are measures of the strength of these laws) that are significant in all analyses but the results on sanctions and the 'consequential' accountability (DJI index) are either weak or not significant. Although previous literature assigns a higher grade to states that adopt sanctions, very few states actually sanction schools. Goertz and Duffy (2003) find that even the few states that practice takeover and reconstitution of failing schools struggle to find financial resources to support such changes. Therefore, only a small proportion of reconstitution eligible schools are ever sanctioned. Our results indicate that strong sanctions do not affect school behavior as much as moderate accountability measures since states lack capacity to implement them.

This study documents the unintended consequences of school accountability laws on medical diagnosis of children and their treatment options. While some schools may be inappropriately labeling marginal students as ADHD, it may also be that children who were previously undiagnosed and went undetected in over-crowded schools, are now detected and treated for their disorder. Nonetheless, it appears that at least some policy makers take the first view: increasingly state and federal laws are being enacted that prevent teachers and other school personnel from requiring the use of psychostimulant drugs for any student, especially as a precondition for attending classes. These laws are motivated in part by the concern that without such protections children are wrongly diagnosed and stigmatized as mentally disabled. However, some states have tightened such laws even further. Connecticut passed a law in 2001 (AB 5701) prohibiting school personnel from even *recommending* the use of psychostimulants to parents for any child, which was followed by similar laws in Illinois and Virginia in 2002 (SB 1719 and HB 90 respectively).

The intended and unintended consequences of school accountability have been actively studied in recent years, and a number of studies have demonstrated that school accountability laws are associated both with improved student outcomes as well as with a number of behavioral consequences. Among the more frequently studied accountability consequences is the effect of accountability on the classification of students with disabilities. Several studies have found that schools subject to more stringent accountability pressure tend to increase their rate of low-achieving students being served by special education services, leading observers to either argue that this is evidence of "gaming of the system" or alternatively of a role for accountability to help to identify students in need of services.

The present study expands upon this literature by investigating whether school accountability leads to medical diagnoses and medication therapy. This is important because these diagnoses require both school decision-making but also the professional judgment of physicians. While physician decisions on ADHD cases are certainly affected by input from school personnel, it provides an interesting opportunity for interpreting whether the behavioral consequences of accountability have been positive or negative, at least along this one dimension. Our first set of results on the probability of diagnosis shows that in the presence of more stringent accountability laws, more children are likely to be diagnosed with ADHD. This result alone could be interpreted as a positive or negative effect of the laws since physicians' decisions are subject to input from school personnel. However, as previously noted (see page 5), drug therapy for ADHD children improves behavior as well as academic achievement and the academic achievement of their peers. Thus, there may be an incentive (among school personnel) to also influence the decision to initiate medication therapy among those who are diagnosed with ADHD but are not on medication therapy. Yet, conditional on diagnosis, our second set of results show that presence of more stringent accountability laws do not change the probability of receiving medication therapy. If physicians were simply reflecting the pressures felt by schools due to the accountability laws and not using independent judgment about diagnosis, we should have seen evidence of more ADHD children being prescribed medication therapy as well. These two results together *suggest* a positive, albeit unintended, consequence of the accountability laws, i.e., children who were previously undiagnosed and went undetected in over-crowded schools, are now detected and treated for their disorder.

APPENDIX A. ACCOUNTABILITY INDEXES

Dee Jacob Index (DJI): The 'consequential accountability' index (DJI in the paper) is based on previous research by Hanushek and Raymond (2005) and Dee and Jacob (2009). To construct this index, we used the information on the year a state became a consequential accountability (CA) state as given in Dee and Jacob (2009, Table 2). However, this information is available for only thirty states prior to NCLB. For the remaining twenty states and D.C., we identified the year it became a consequential accountability state as the first year when it adopted either assistance, rewards or sanctions (as identified by Quality Counts) and is coded as a CA state for all periods after that. Further, all states are coded as one in 2003 due to the NCLB act.

Carnoy Leob Index (CLI): Accountability pressures index developed by Carnoy and Loeb (2002) was reconstructed using Quality Counts data for all years used in our analysis. The index coding follows the following rules in Carnoy and Loeb (2002, pp.311):

> States receiving a zero do not test students statewide or do not set any statewide standards for schools or districts. States that require state testing in the elementary and middle grades and the reporting of test results to the state but no school (or district) sanctions or rewards (no or weak external pressure) get a 1. Those states that test at the elementary and middle school levels and have moderate school or district accountability sanctions / rewards or, alternatively, a high school exit test (that sanctions students but pressures schools to improve student performance) get a 2. Those states that test at the lower and middle grades, have moderate accountability repercussions for schools and districts, and require an exit test in high school, get a 3. Those that test and place strong pressure on schools or districts to improve student achievement (threat of reconstitution, principal transfer, loss of students) but do not require a high school exit test get a 4. States receiving a 5 test students in primary and middle grades, strongly sanction and reward schools or districts based on improvement in student test scores, and require a high school minimum competency exit test for graduation.

Table below shows Carnoy and Loeb criteria and Quality counts variable that we used to construct index above.

Carnoy and Loeb Criteria	Quality Counts Variable
State developed statewide standards for schools	Has the state adopted standards in core academic subjects?
State requires statewide testing	State plans to participate in NAEP Or state-wide testing in english and math at elementary and middle school levels
States test at elementary and middle school levels	State-wide testing in english and math at elementary and middle school levels
State requires a high school minimum competency exit test for graduation	Graduation contingent upon performance on state-wide exit or end-of-course exam
Monetary rewards	State provides rewards to high performing or improved schools: includes monetary rewards that can be used for bonuses or school improvement
Moderate school accountability sanctions: write school improvement plan, watch lists, warnings	Schools named low-performing receive assistance: on-site external team provides assistance, school improvement plan, adoption of research-based program, professional development
State puts a strong pressure on schools or districts to improve student achievement (threat of reconstitution, principal transfer, loss of students)	State accountability system includes following sanctions: closures, student transfer, reconstitution, turning schools to private management, withholding funds

In addition, some states do not fall into any of the categories described by Carnoy and Loeb (2002). For example, a number of states conduct state-wide testing but do not adopt report cards. For such states a grade of 0.5 was assigned. In addition, some states do not use report cards or rate schools based on academic performance but have a threat of consequences (e.g. Utah does not use report cards or ratings but adopted monetary rewards for schools). Consequences in such states do not make much sense and such states were rated as having an index value of one.

Appendix B. Detailed Regression Results

Table B-1. Regression of Log Quantity on State Accountability Laws

		Dependent Variable: Ln Qnty ADHD drugs										Dependent Variable: Ln Qnty non-ADHD drugs	
(N = 240)	Mean (Std)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Report Cards Law	0.888	0.00052											
	(0.317)	(0.016)											
Rates Law	0.638		0.017										
	(0.482)		(0.013)										
Assistance Law	0.567			0.042^{a}			0.038^{a}	0.035^{b}					
	(0.497)			(0.012)			(0.012)	(0.014)					
Rewards Law	0.350				0.032^{b}		0.024^{c}	0.024^{c}					
	(0.478)				(0.015)		(0.014)	(0.014)					
Sanctions Law	0.417					0.020		0.0053					
	(0.494)					(0.014)		(0.017)					
QCI: Quality Counts Index	1.333								0.021^{a}			-0.014	
(Sum of the three laws)	(1.134)								(0.0067)			(0.021)	
CLI: Carnoy-Loeb Index	2.710									0.0070^{c}			-0.013
	(0.688)									(0.0038)			(0.016)
DJI: Dee-Jacob Index	0.742										-0.0086		
	(.439)										(0.016)		
Year 2000	0.200	0.19^{a}	0.19^{a}	0.19^{a}	0.20^{a}	0.19^{a}	0.19^{a}	0.19^{a}	0.19^{a}	0.19^{a}	0.19^{a}	0.56^{a}	0.56^{a}
	(0.401)	(0.034)	(0.035)	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)	(0.033)	(0.035)	(0.033)	(0.088)	(0.088)
Year 2001	0.200	0.28^{a}	0.28^{a}	0.27^{a}	0.28^{a}	0.27^{a}	0.27^{a}	0.27^{a}	0.27^{a}	0.27^{a}	0.28^{a}	0.82^{a}	0.82^{a}
	(0.401)	(0.039)	(0.041)	(0.042)	(0.039)	(0.039)	(0.042)	(0.041)	(0.040)	(0.040)	(0.039)	(0.12)	(0.12)
Year 2002	0.200	0.43^{a}	0.43^{a}	0.42^{a}	0.43^{a}	0.42^{a}	0.43^{a}	0.42^{a}	0.42^{a}	0.43^{a}	0.43^{a}	1.21^{a}	1.21^{a}
	(0.401)	(0.049)	(0.051)	(0.052)	(0.049)	(0.050)	(0.052)	(0.052)	(0.051)	(0.050)	(0.048)	(0.15)	(0.16)
Year 2003	0.200	0.71^{a}	0.70^{a}	0.69^{a}	0.71^{a}	0.69^{a}	0.69^{a}	0.69^{a}	0.69^{a}	0.69^{a}	0.71^{a}	1.46^{a}	1.46^{a}

Notes: All regressions include state dummies. The sample consists of 48 States (Hawaii, Alaska and D.C. are excluded). Standard errors (clustered by state) are in parenthesis and a, b, c are significance levels at 1,5 and 10% respectively.

Table B-1 – continued from previous page

(N = 240)	Mean (Std)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		Dependent Variable: Ln Qnty ADHD drugs										Dependent Variable: Ln Qnty non-ADHD drugs	
	(0.401)	(0.056)	(0.059)	(0.060)	(0.057)	(0.059)	(0.061)	(0.062)	(0.060)	(0.058)	(0.056)	(0.19)	(0.19)
Ln Population	15.122	-0.37	-0.40	-0.35	-0.41	-0.42	-0.39	-0.40	-0.45	-0.47	-0.34	3.06^{b}	3.19^{b}
	(0.992)	(0.44)	(0.45)	(0.45)	(0.47)	(0.45)	(0.47)	(0.48)	(0.47)	(0.45)	(0.44)	(1.51)	(1.55)
% Black Pop	10.518	0.046	0.045	0.048	0.032	0.048	0.038	0.038	0.040	0.041	0.047	0.20^{c}	0.21^{b}
	(9.632)	(0.043)	(0.042)	(0.040)	(0.045)	(0.043)	(0.042)	(0.042)	(0.041)	(0.041)	(0.043)	(0.10)	(0.10)
% Other Pop	4.865	-0.055^{a}	-0.053^{a}	-0.052^{a}	-0.057^{a}	-0.052^{a}	-0.054^{a}	-0.053^{a}	-0.052^{a}	-0.051^{a}	-0.055^{a}	-0.066	-0.071
	(3.192)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)	(0.061)	(0.061)
% Age 5-19	21.524	0.042^{b}	0.044^{b}	0.037^{c}	0.040^{b}	0.039^{c}	0.037^{c}	0.036^{c}	0.035^{c}	0.036^{c}	0.041^{b}	0.054	0.061
	(1.293)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.077)	(0.079)
Ln Special Ed	11.316	0.25	0.22	0.21	0.22	0.22	0.19	0.18	0.17	0.24	0.26	-1.10	-1.14^{c}
(Age 3-17)	(0.967)	(0.33)	(0.32)	(0.31)	(0.32)	(0.32)	(0.30)	(0.30)	(0.30)	(0.32)	(0.33)	(0.68)	(0.68)
Ln School Lunch	12.804	-0.27	-0.26	-0.32	-0.22	-0.21	-0.27	-0.25	-0.19	-0.22	-0.29	-0.31	-0.35
Participation	(0.983)	(0.24)	(0.24)	(0.22)	(0.23)	(0.24)	(0.21)	(0.21)	(0.23)	(0.24)	(0.25)	(0.79)	(0.81)
Pupil Teacher Ratio	15.520	-0.0040	-0.0054	-0.0081	-0.0039	-0.0062	-0.0076	-0.0080	-0.0085	-0.0046	-0.0038	0.0024	0.00047
	(2.286)	(0.010)	(0.0096)	(0.0087)	(0.010)	(0.010)	(0.0091)	(0.0090)	(0.010)	(0.011)	(0.0100)	(0.033)	(0.033)
Unemployment Rate	4.617	0.0022	0.0027	0.0022	0.0041	0.0029	0.0037	0.0039	0.0043	0.0011	0.0017	-0.0025	0.00093
	(1.152)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.051)	(0.051)
Log Percapita Income	10.235	0.19	0.14	0.046	0.11	0.22	0.0027	0.018	0.095	0.14	0.19	0.33	0.35
	(0.145)	(0.36)	(0.36)	(0.34)	(0.35)	(0.36)	(0.33)	(0.33)	(0.34)	(0.36)	(0.36)	(1.33)	(1.32)
Percent Uninsured	13.740	0.0076	0.0073	0.0076	0.0083	0.0077	0.0081^{c}	0.0082^{c}	0.0082^{c}	0.0078	0.0077	0.015	0.015
	(3.872)	(0.0053)	(0.0051)	(0.0048)	(0.0051)	(0.0052)	(0.0047)	(0.0048)	(0.0048)	(0.0052)	(0.0053)	(0.015)	(0.015)
Ln SCHIP $	15.479	0.0043^{c}	0.0040	0.0037	0.0045^{c}	0.0038	0.0039	0.0038	0.0036	0.0039	0.0044^{c}	0.0052	0.0056
(State Component)	(2.858)	(0.0024)	(0.0026)	(0.0026)	(0.0025)	(0.0025)	(0.0026)	(0.0027)	(0.0025)	(0.0026)	(0.0024)	(0.0091)	(0.0090)
Ln Medicaid	5.840	0.037	0.042	0.045	0.046	0.040	0.051	0.051	0.050	0.043	0.036	-0.066	-0.068
Population	(1.049)	(0.039)	(0.037)	(0.035)	(0.036)	(0.039)	(0.035)	(0.035)	(0.035)	(0.037)	(0.038)	(0.14)	(0.14)

Notes: All regressions include state dummies. The sample consists of 48 States (Hawaii, Alaska and D.C. are excluded). Standard errors (clustered by state) are in parenthesis and a, b, c are significance levels at 1,5 and 10% respectively.

Table B-2. Probability of Diagnosis

	Main Analysis: Public School Students (1-6)						Double Falsification Test (7-10)			
							Pvt. Sch. Students		Pub. Sch. Students 1000 Bootstraps on	
	N=49,527						N=6,714		N=6,714	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Assistance Law	0.13^{a}			0.15^{a}						
(1/0: 1 if State Law in effect)	(0.028)			(0.030)						
Rewards Law		0.076^{b}		0.079^{a}						
(1/0: 1 if State Law in effect)		(0.038)		(0.031)						
Sanctions Law			0.036	-0.046						
(1/0: 1 if State Law in effect)			(0.041)	(0.030)						
QCI: Quality Counts Index					0.055^{a}		-0.021		0.058^{c}	
(Sum of the three laws)					(0.017)		(0.037)		(0.040)	
CLI: Carnoy-Loeb Index						0.034^{a}		-0.017		0.036^{c}
						(0.011)		(0.024)		(0.029)
Age	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.015^{a}	0.056^{a}	0.056^{a}	0.015^{c}	0.015^{c}
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.010)	(0.010)	(0.011)	(0.011)
Gender (1/0)	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.527^{a}	0.526^{a}
(1 if Male)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.079)	(0.079)	(0.079)	(0.079)
Race (1/0)	-0.30^{a}	-0.30^{a}	-0.29^{a}	-0.31^{a}	-0.30^{a}	-0.30^{a}	-0.053	-0.052	-0.320^{a}	-0.317^{a}
(1 if African American)	(0.049)	(0.053)	(0.050)	(0.051)	(0.051)	(0.050)	(0.20)	(0.20)	(0.121)	(0.121)
Race (1/0)	-0.17^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.16^{a}	-0.26^{a}	-0.26^{a}	-0.189	-0.191
(1 if Other)	(0.051)	(0.052)	(0.052)	(0.051)	(0.051)	(0.051)	(0.099)	(0.098)	(0.168)	(0.168)
Ethnicity (1/0)	-0.37^{a}	-0.36^{a}	-0.37^{a}	-0.36^{a}	-0.36^{a}	-0.37^{a}	-0.10	-0.10	-0.381^{b}	-0.385^{b}
(1 if Hispanic)	(0.074)	(0.075)	(0.074)	(0.075)	(0.075)	(0.075)	(0.13)	(0.13)	(0.207)	(0.207)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: Columns (7-10) show the mean and standard deviation of the 1000 bootstraps (with replacement). The bootstraps adjust for the observation weights in the data set.

Table B-2 – continued from previous page

	Main Analysis: Public School Students (1-6)						Double Falsification Test (7-10)			
							Pvt. Sch. Students		Pub. Sch. Students 1000 Bootstraps on	
	N=49,527						N=6,714		N=6,714	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Family Structure 1 (1/0)	0.33^a	0.33^a	0.33^a	0.33^a	0.33^a	0.33^a	0.46^a	0.46^a	0.328^a	0.328^a
(1 if Step Family, 2 Parents)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.11)	(0.11)	(0.107)	(0.107)
Family Structure 2 (1/0)	0.26^a	0.26^a	0.26^a	0.26^a	0.26^a	0.26^a	0.32^a	0.32^a	0.268^a	0.268^a
(1 if Single Mother, No Father)	(0.040)	(0.039)	(0.039)	(0.040)	(0.040)	(0.040)	(0.086)	(0.086)	(0.102)	(0.102)
Family Structure 3 (1/0)	0.17^a	0.17^a	0.17^a	0.17^a	0.17^a	0.17^a	0.13	0.13	0.164	0.164
(1 if Other)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.19)	(0.19)	(0.153)	(0.153)
Poverty Level 1 (1/0)	0.24^a	0.24^a	0.24^a	0.24^a	0.24^a	0.24^a	-0.022	-0.021	0.238^a	0.239^a
(1 if < 100% of Poverty Line (P.L.))	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)	(0.17)	(0.18)	(0.163)	(0.163)
Poverty Level 2 (1/0)	0.12^b	0.12^b	0.12^b	0.12^b	0.12^b	0.12^b	0.10	0.10	0.112	0.113
(1 if 100% to <133% P.L.)	(0.053)	(0.053)	(0.054)	(0.053)	(0.053)	(0.054)	(0.23)	(0.23)	(0.174)	(0.174)
Poverty Level 3 (1/0)	0.069	0.063	0.066	0.067	0.067	0.068	0.028	0.027	0.034	0.034
(1 if 133% to <150% P.L.)	(0.076)	(0.078)	(0.077)	(0.076)	(0.077)	(0.076)	(0.38)	(0.38)	(0.235)	(0.235)
Poverty Level 4 (1/0)	0.049	0.050	0.049	0.050	0.050	0.050	0.14	0.14	0.049	0.049
(1 if 150% to <185% P.L.)	(0.039)	(0.040)	(0.039)	(0.040)	(0.040)	(0.040)	(0.14)	(0.14)	(0.166)	(0.166)
Poverty Level 5 (1/0)	0.057	0.055	0.055	0.057	0.055	0.054	0.22	0.22	0.022	0.021
(1 if 185% to <200% P.L.)	(0.055)	(0.054)	(0.055)	(0.054)	(0.054)	(0.054)	(0.23)	(0.23)	(0.216)	(0.216)
Poverty Level 6 (1/0)	0.074^b	0.073^b	0.073^b	0.074^b	0.073^b	0.073^b	0.093	0.093	0.069	0.069
(1 if 200% to <300% P.L.)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.26)	(0.26)	(0.113)	(0.113)
Poverty Level 7 (1/0)	-0.080^c	-0.079	-0.080^c	-0.079^c	-0.080^c	-0.080^c	-0.088	-0.089	-0.080	-0.080
(1 if 300% to <400% P.L.)	(0.047)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.081)	(0.081)	(0.112)	(0.112)
Highest Education in Household	-0.023	-0.025	-0.022	-0.027	-0.026	-0.024	0.30	0.30	-0.056	-0.054
(1/0: 1 if Less than High School)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.25)	(0.25)	(0.236)	(0.236)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: Columns (7-10) show the mean and standard deviation of the 1000 bootstraps (with replacement). The bootstraps adjust for the observation weights in the data set.

Table B-2 – continued from previous page

	Main Analysis: Public School Students (1-6)						Double Falsification Test (7-10)			
							Pvt. Sch. Students		Pub. Sch. Students 1000 Bootstraps on	
	N=49,527						N=6,714		N=6,714	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Highest Education in Household	-0.012	-0.015	-0.012	-0.015	-0.015	-0.014	0.19^{c}	0.19^{c}	-0.015	-0.014
(1/0: 1 if High School)	(0.035)	(0.036)	(0.035)	(0.036)	(0.035)	(0.035)	(0.11)	(0.11)	(0.101)	(0.101)
% Age 5-19	-0.021	-0.0045	-0.020	-0.0050	-0.0085	-0.018	0.055^{c}	0.058^{b}	-0.008	-0.018
(State-level Var)	(0.015)	(0.016)	(0.018)	(0.014)	(0.015)	(0.015)	(0.028)	(0.027)	(0.041)	(0.042)
Ln Special Ed (Age 3-17)	-0.065	-0.053	-0.11	-0.0012	-0.040	-0.072	0.077	0.083	-0.046	-0.080
(State-level Var)	(0.094)	(0.093)	(0.091)	(0.093)	(0.099)	(0.089)	(0.23)	(0.23)	(0.215)	(0.211)
Ln School Lunch Participation	0.20^{a}	0.12	0.21^{a}	0.10	0.14^{c}	0.18^{b}	-0.27	-0.28	0.144	0.183
(State-level Var)	(0.071)	(0.082)	(0.079)	(0.068)	(0.076)	(0.073)	(0.18)	(0.18)	(0.204)	(0.205)
Pupil Teacher Ratio	-0.0030	-0.0081	-0.0048	-0.0063	-0.0079	-0.0014	-0.013	-0.016	-0.009	-0.003
(State-level Var)	(0.0076)	(0.0067)	(0.0081)	(0.0061)	(0.0067)	(0.0078)	(0.017)	(0.017)	(0.020)	(0.020)
Ln SCHIP Compensation	0.0020	0.0041	-0.0017	0.0085	0.016	0.0060	-0.012	-0.010	0.017	0.007
(State-level Var)	(0.0084)	(0.012)	(0.015)	(0.0097)	(0.013)	(0.014)	(0.022)	(0.020)	(0.037)	(0.035)
Ln Medicaid Population	-0.12	-0.039	-0.086	-0.076	-0.080	-0.10	0.24^{c}	0.26^{c}	-0.083	-0.107
(State-level Var)	(0.077)	(0.069)	(0.079)	(0.064)	(0.076)	(0.079)	(0.15)	(0.15)	(0.154)	(0.16)
Ln SCHIP Population	-0.0049	-0.0073	0.00048	-0.014	-0.024	-0.010	0.0100	0.0066	-0.025	-0.010
(State-level Var)	(0.014)	(0.019)	(0.023)	(0.016)	(0.020)	(0.021)	(0.033)	(0.030)	(0.054)	(0.051)
Constant	-2.60^{a}	-2.44^{a}	-2.41^{a}	-2.68^{a}	-2.59^{a}	-2.49^{a}	-2.22^{c}	-2.24^{c}	-2.557^{b}	-2.452^{b}
	(0.53)	(0.51)	(0.54)	(0.51)	(0.56)	(0.56)	(1.27)	(1.25)	(1.200)	(1.189)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: Columns (7-10) show the mean and standard deviation of the 1000 bootstraps (with replacement). The bootstraps adjust for the observation weights in the data set.

TABLE B-3. Probability of Medication

	Pr(MEDS=1)		Pr(MEDS=1|ADHD=1)	
	(1)	(2)	(3)	(4)
QCI: Quality Counts Index	0.051^{b}		0.096	
(Sum of the three laws)	(0.021)		(0.077)	
CLI: Carnoy-Loeb Index		0.033^{b}		0.053
		(0.014)		(0.060)
$\widehat{\widehat{D}}$: Probability of ADHD			5.69^{c}	5.48
			(3.14)	(3.49)
$\widehat{\widehat{D}}\times$ Index			-1.02^{b}	-0.53
			(0.49)	(0.35)
Age	-0.013^{a}	-0.013^{a}	-0.11^{a}	-0.11^{a}
	(0.0043)	(0.0043)	(0.011)	(0.012)
Gender	0.47^{a}	0.47^{a}	-0.26	-0.24
(1/0: 1 if Male)	(0.032)	(0.032)	(0.25)	(0.25)
Race	-0.35^{a}	-0.35^{a}	-0.087	-0.10
(1/0: 1 if African American)	(0.056)	(0.057)	(0.15)	(0.15)
Race	-0.18^{a}	-0.18^{a}	-0.0056	-0.011
(1/0: 1 if Other)	(0.063)	(0.063)	(0.14)	(0.14)
Ethnicity	-0.43^{a}	-0.43^{a}	-0.060	-0.078
(1/0: 1 if Hispanic)	(0.049)	(0.048)	(0.20)	(0.20)
Family Structure	0.23^{a}	0.24^{a}	-0.40^{c}	-0.39^{c}
(1/0: 1 if Step Family, 2 Parents)	(0.045)	(0.045)	(0.21)	(0.21)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: In column (3), $\widehat{\widehat{D}}$ and $\widehat{\widehat{D}}\times$ Index is computed using the Quality Counts index. Similarly, in columns (4) these two variables are computed using the Carnoy-Loeb index.

Table B-3 – continued from previous page

	Pr(MEDS=1)		Pr(MEDS=1|ADHD=1)	
	(1)	(2)	(3)	(4)
Family Structure (1/0: 1 if Single Mother, No Father)	0.26^a (0.048)	0.26^a (0.048)	-0.13 (0.15)	-0.11 (0.15)
Family Structure (1/0: 1 if Other)	0.068 (0.065)	0.069 (0.065)	-0.36^b (0.16)	-0.35^b (0.16)
Poverty Level (1 if < 100% of Poverty Line (P.L.))	0.075 (0.068)	0.076 (0.068)	-0.59^a (0.15)	-0.58^a (0.16)
Poverty Level (1/0: 1 if 100% to <133% P.L.)	0.087 (0.067)	0.088 (0.067)	-0.21 (0.20)	-0.20 (0.20)
Poverty Level (1/0: 1 if 133% to <150% P.L.)	-0.071 (0.085)	-0.071 (0.085)	-0.42^c (0.21)	-0.41^c (0.21)
Poverty Level (1/0: 1 if 150% to <185% P.L.)	0.074 (0.059)	0.074 (0.059)	-0.027 (0.23)	-0.023 (0.23)
Poverty Level (1/0: 1 if 185% to <200% P.L.)	0.0075 (0.084)	0.0069 (0.084)	-0.17 (0.19)	-0.16 (0.19)
Poverty Level (1/0: 1 if 200% to <300% P.L.)	-0.013 (0.059)	-0.013 (0.059)	-0.32^a (0.12)	-0.31^b (0.12)
Poverty Level (1/0: 1 if 300% to <400% P.L.)	-0.078^b (0.039)	-0.078^b (0.039)	0.037 (0.074)	0.029 (0.077)
Highest Education in Household (1/0: 1 if Less than High School)	0.048 (0.13)	0.050 (0.13)	0.16 (0.22)	0.15 (0.22)
Highest Education in Household (1/0: 1 if High School)	-0.0088 (0.034)	-0.0080 (0.034)	-0.035 (0.061)	-0.037 (0.061)

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: In column (3), $\widehat{\widehat{D}}$ and $\widehat{\widehat{D}}\times$ Index is computed using the Quality Counts index. Similarly, in columns (4) these two variables are computed using the Carnoy-Loeb index.

Table B-3 – continued from previous page

	Pr(MEDS=1)		Pr(MEDS=1|ADHD=1)	
	(1)	(2)	(3)	(4)
% Age 5-19	0.00030	-0.0084	0.018	0.023
(State-level Var)	(0.022)	(0.022)	(0.038)	(0.037)
Ln Special Ed (Age 3-17)	-0.16	-0.18	-0.33	-0.32
(State-level Var)	(0.12)	(0.12)	(0.25)	(0.26)
Ln School Lunch Participation	0.27^{b}	0.30^{a}	0.30	0.28
(State-level Var)	(0.11)	(0.11)	(0.22)	(0.23)
Pupil Teacher Ratio	-0.022^{b}	-0.016	-0.043^{b}	-0.047^{b}
(State-level Var)	(0.011)	(0.010)	(0.021)	(0.021)
Ln SCHIP	0.026	0.017	0.028	0.037
Compensation (State Component)	(0.018)	(0.018)	(0.034)	(0.031)
Ln Medicaid Population	-0.092	-0.11	0.027	0.044
	(0.080)	(0.085)	(0.11)	(0.11)
Ln S-CHIP Population	-0.037	-0.023	-0.039	-0.052
	(0.028)	(0.027)	(0.053)	(0.049)
Constant	-2.71^{a}	-2.61^{a}	1.48	1.41
	(0.66)	(0.65)	(1.10)	(1.08)
Observations	49513	49513	4701	4701

Note 1: All specifications restrict the sample to children age 5-17 and excludes home-schooled and not in school children.
Note 2: Point estimates adjust for the weights provided in the data set. Standard errors are clustered by states and are in parentheses. Significance levels: (a,b,c) are 1,5 and 10% respectively.
Note 3: In column (3), $\widehat{\widehat{D}}$ and $\widehat{\widehat{D}}\times$ Index is computed using the Quality Counts index. Similarly, in columns (4) these two variables are computed using the Carnoy-Loeb index.

References

Aizer, Anna, "Peer effects and human capital accumulation: The externalities of ADD," NBER Working Paper 14354, National Bureau of Economic Research September 2008.

Aleman, S. R., "Special education for children with attention deficit disorder: Current issues," *Congressional Research Service*, 1991.

American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4 ed. 1994. 4th. Edition.

Anderson, Patricia M. and Kristin F. Butcher, "Reading, writing, and refreshments: Do school finances contribute to childhood obesity?," *Journal of Human Resources*, Summer 2006, *41* (3), 467–494.

Ansell, Susan E., "Put to the test," *Education Week*, January 2004, *23* (17), 75–6.

Arcia, Emily, Maria C. Fernandez, Marisela Jaquez, Hector Castillo, and Maria Ruiz, "Modes of entry into services for young children with disruptive behaviors," *Qualitative Health Research*, November 2004, *14* (9), 1211–1226.

Barbaresi, William J., Slavica K. Katusic, Robert C. Colligan, Amy L. Weaver, Cynthia L. Leibson, and Steven J. Jacobsen, "Long-term stimulant medication treatment of attention-deficit/hyperactivity disorder: Results from a population-based study," *Journal of Developmental & Behavioral Pediatrics*, February 2006, *27* (1), 1–10.

Biederman, Joseph and Stephen V. Faraone, "The effects of attention-deficit/hyperactivity disorder on employment and household income," *Medscape General Medicine*, 2006, *8* (3), 12.

Bokhari, Farasat A.S. and Gary M. Fournier, "Entry in the ADHD drugs market: Welfare impact of generics and me-toos," May 2010.

———, **Frank Heiland, Peter Levine, and G. Thomas Ray**, "Risk factors for discontinuing drug therapy among children with ADHD," *Health Services and Outcomes Research Methodology*, September 2008, *8* (3), 134–158.

Browder, Diane M. and Karena Cooper-Duffy, "Evidence-based practices for students with severe disabilities and the requirement of accountability in "No Child Left Behind"," *Journal of Special Education*, November 2003, *37* (3), 157–163.

Brown, R.T., R.W. Amler, W.S. Freeman, J.M. Perrin, M.T. Stein, H.M Feldman, K. Pierce, M.L. Wolraich, and American Academy of Pediatrics - Committee on Quality Improvement, Subcommittee on Attention-Deficit/Hyperactivity Disorder, "Treatment of Attention-Deficit/Hyperactivity Disorder: Overview of the evidence," *Pediatrics*, June 2005, *115* (6), e749–e757.

Carlson, Caryn L., William E. Pelham, Richard Milich, and Joanne Dixon, "Single and combined effects of methylphenidate and behavior therapy on the classroom performance of children with attention-deficit hyperactivity disorder," *Journal of Abnormal Child Psychology*, April 1992, *20* (2), 213–32.

Carnoy, Martin and Susanna Loeb, "Does external accountability affect student outcomes? A cross-state analysis," *Educational Evaluation and Policy Analysis*, December 2002, *24* (2), 305–331.

Centers for Disease Control and Prevention, "Mental health in the United States: Prevalence of diagnosis and medication treatment for attention-deficit/hyperactivity disorder –United States, 2003.," *Morbidity and Mortality Weekly Report*, September 2 2005, *54* (34), 842–847.

Chan, Eugenia, Marianne R. Hopkins, James M. Perrin, Carla Herrerias, and Charles J. Homer, "Diagnostic practices for attention deficit hyperactivity disorder: A national survey of primary care physicians," *Ambulatory Pediatrics*, July-August 2005, *5* (4), 201–208.

Cullen, Julie Berry and Randall Reback, "Tinkering toward accolades: School gaming under a performance accountability system," NBER Working Paper 12286, National Bureau of Economic Research June 2006.

Dee, Thomas and Brian Jacob, "The impact of no child left behind on student achievement," NBER Working Paper 15531, National Bureau of Economic Research November 2009.

Evans, Steven W. and William E. Pelham, "Psychostimulant effects on academic and behavioral measures for ADHD junior high school students in a lecture format classroom," *Journal of Abnormal Child Psychology*, October 1991, *19* (5), 537–52.

Figlio, David N., "Testing, crime and punishment," *Journal of Public Economics*, 2006, *90*, 837–851.

——— **and Joshua Winicki**, "Food for thought: the effects of school accountability plans on school nutrition," *Journal of Public Economics*, February 2005, *89* (2-3), 381–394.

——— **and Lawrence S. Getzler**, "Accountability, ability and disability: Gaming the system?," in Gronberg Timothy J. and Dennis W. Jansen, eds., *Improving School Accountability - Check-Ups or Choice*, Vol. 14 of *Advances in Applied Microeconomics*, Elsevier, 2006, pp. 35–49.

Finn, Chester E. and Marci Kanstoroom, "State academic standards," in D. Ravitch, ed., *Brookings Papers on Education Policy*, Brookings Institution Press, 2001, pp. 131–179.

Goertz, Margaret and Mark Duffy, "Mapping the landscape of high-stakes testing and accountability programs," *Theory Into Practice*, Winter 2003, *42* (1), 4–11.

Greene, Ross W., Sara K. Beszterczey, Tai Katzenstein, Kenneth Park, and Jennifer Goring, "Are students with ADHD more stressful to teach? Patterns of teacher stress in an elementary school sample," *Journal of Emotional and Behavioral Disorders*, Summer 2002, *10* (2), 79–89.

Hanushek, Eric A. and Margaret F. Raymond, "Does school accountability lead to improved student performance?," *Journal of Policy Analysis and Management*, Spring 2005, *24* (2), 297–327.

Jacob, Brian A., "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools," *Journal of Public Economics*, June 2005, *89* (5-6), 761–796.

Koretz, Daniel M., "Arriving in Lake Wobegon: Are standardized tests exaggerating achievement and distorting instruction?," *American Educator*, 1988, *12* (2), 8–15,46–52.

______, "Limitations in the use of achievement tests as measures of educators' productivity," *Journal of Human Resources*, Autumn 2002, *37* (4), 752–777.

______ **and Sheila I. Barron**, "The validity of gains in scores on the Kentucky instructional results information system (KIRIS)," Rand Corp., Santa Monica, CA. Inst. for Education and Training 1998.

Langberg, Joshua M., Tanya E. Froehlich, Richard E.A. Loren, Jessica E. Martin, and Jeffery N. Epstein, "Assessing children with ADHD in primary care settings," *Expert Review of Neurotherapeutics*, April 2008, *8* (4), 627–641.

Lazear, Edward P., "Educational production," *The Quarterly Journal of Economics*, August 2001, *CXVI* (3), 777–803.

Lee, Jaekyung and Kenneth K. Wong, "The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes," *American Educational Research Journal*, December 2004, *41* (4), 797–832.

Maddala, G.S., *Limited dependent and qualitative variables in econometrics*, Vol. 3 of *Econometric Society Monographs*, New York: Cambridge University Press, 1983.

Mannuzza, S., R.G. Klein, A. Bessler, P. Malloy, and M.E. Hynes, "Educational and occupational outcome of hyperactive boys grown up," *Journal of the American Academy of Child and Adolescent Psychiatry*, September 1997, *36* (9), 1222–1227.

MTA Cooperative Group, Stephen P. Hinshaw, Glen R. Elliott, C. Keith Conncers, Karen C. Wells, John March, James Swanson, Timothy Wigal, Dennis P. Cantwell, H.B. Abikoff, L. Hecthman, L.L. Greenhill, and J.H. Newcorn, "A 14-month randomized clinical trial of treatment strategies for ADHD," *Archives of General Psychiatry*, December 1999, *56*, 1073–1086.

______, ______, ______, ______, ______, ______, ______, ______, ______, ______, ______, ______, **and** ______, "Moderators and mediators of treatment response for children with ADHD: The MTA Study," *Archives of General Psychiatry*, 1999, *56*, 1088–1096.

Murphy, Kevin and Russell A. Barkley, "Attention deficit hyperactivity disorder adults: comorbidities and adaptive impairments," *Comprehensive Psychiatry*, November-December 1996, *37* (6), 393–401.

Neal, Derek and Diane Whitmore Schanzenbach, "Left behind by design: Proficiency counts and test-based accountability," NBER Working Paper 13293, National Bureau of Economic Research August 2007.

Olfson, Mark, Steven C. Marcus, M.M. Weissman, and Peter S. Jensen, "National trends in the use of psychotropic medications by children," *Journal of the American Academy of Child and Adolescent Psychiatry*, May 2002, *41* (5), 514–521.

Pelham, William E., Elizabeth M. Gnagy, Andrew R. Greiner, Betsy Hoza, Stephen P. Hinshaw, James M. Swanson, Steve Simpson, Cheri Shapiro, Oscar Bukstein, Carrie Baron-Myak, and Keith McBurnett, "Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer treatment program," *Journal of Abnormal Child Psychology*, December 2000, *28* (6), 507–525.

Reback, Randall, "Teaching to the rating: School accountability and the distribution of student achievement," *Journal of Public Economics*, June 2008, *92* (5-6), 1394–1415.

Safer, Daniel J. and John M. Krager, "A survey of medication treatment for hyperactive/inattentive students," *Journal of the American Medical Association*, October 1988, *260* (15), 2256–2258.

———, **Julie Magno Zito, and Eric M. Fine**, "Increased methylphenidate usage for attention deficit disorder in the 1990s," *Pediatrics*, December 1996, *98*, 1984–1988.

Sax, Leonard and Kathleen J. Kautz, "Who first suggests the diagnosis of attention-deficit/hyperactivity disorder," *Annales of Family Medicine*, Sep-Oct 2003, *1* (3), 171–174.

Scheffler, Richard M., Timothy T. Brown, Brent D. Fulton, Stephen P. Hinshaw, Peter Levine, and Susan Stone, "Positive association between attention-deficit/hyperactivity disorder medication use and academic achievement during elementary school," *Pediatrics*, May 2009, *123* (5), 1273–1279.

Schneider, Helen and Daniel Eisenberg, "Who receives a diagnosis of attention-deficit/ hyperactivity disorder in the United States elementary school population?," *Pediatrics*, April 2006, *117* (4), e601–e609.

Secnik, Kristina, Andrine Swensen, and Maureen J. Lage, "Comorbidities and costs of adult patients diagnosed with attention-deficit hyperactivity disorder," *PharmacoEconomics*, 2005, *23* (1), 93–102.

Smith, Thomas J. and Gail Adams, "The effect of comorbid Ad/HD and learning disabilities on parent-reported behavioral and academic outcomes of children," *Learning Disabilities Quarterly*, Spring 2006, *29* (2), 101–112.

U.S. Department of Education, "A guide to the individualized education program," Technical Report, Office of Special Education and Rehabilitation Services, Washington, D.C. July 2000.

———, "No Child Left Behind: A Desktop Reference," Technical Report, Office of Elementary and Secondary Education, Washington, D.C. 2002.

Wong, Manyee, Thomas D. Cook, and Peter M. Steiner, "No child left behind: An interim evaluation of its effects on learning using two interrupted time series each with its own non-equivalent comparison series," IPR Working Paper 09-11, Institute for Policy Research, Northwestern University 2010.