Online Appendix

Title: State Variation in Increased ADHD Prevalence: Links to NCLB School Accountability and State Medication Laws

Version Date: May 5, 2015

I. Overview

The purpose of this online appendix is to show how a difference-in-differences (DiD) result from a logistic regression model can be used to estimated average marginal effects in more interpretable probability units (1). We derive the average marginal effect of the DiD result in model 4 in the main article. Furthermore, we describe how we report the average marginal effect in Table 4 (main article) and display the average marginal effect in Figure 1 (main article).

II. Transforming Difference-in-Differences Logistic Regression Results into Average Marginal Effects in Probability Units

A differenced model includes a DiD term that compares how the outcome of interest changes over time between observations within two groups (2). The DiD term is aptly named, because it is, indeed, a difference in differences. Difference 1 is the difference in the outcome of interest between two points in time for observations in group 1. Difference 2 is the same difference for observations in group 2. The DiD compares difference 1 with difference 2, by differencing them.

In model 4 in the main article, the DiD term, $Med_s \times Wave_t$, estimates the association between changes in ADHD diagnostic prevalence and psychotropic medication laws. However, the DiD parameter in model 4 represents the ratio of two odds ratios, which is difficult to interpret. We transform the DiD result from the logistic regression model into an average marginal effect in probability units, to aid in the interpretation of the result. In probability units, the DiD average marginal effect's structure is still a difference-in-differences, but instead of representing the ratio of two odds ratios, it represents the difference of two differences, where each difference is defined by the change in two probabilities. The first difference is the change in ADHD diagnostic prevalence between two National Survey of Children's Health (NSCH) waves in states that had a psychotropic medication law that prohibit public schools from recommending or requiring medication use. The second difference is the same change for states without such law. Therefore, the DiD average marginal effect compares ADHD prevalence changes between two NSCH waves for these two groups of states.

II.A. Derivation of the Average Marginal Effect for a Difference-in-Differences Logistic Regression Result

Our statistical analyses that generated the results in Table 2 (main article) are based on the logistic regression model shown in Eq (1). The four sets of results in the table vary by the income-levels of the included children and which two NSCH waves are included.

$$\ln\left(\frac{p(ADHD_{i,s,t}=1)}{1-p(ADHD_{i,s,t}=1)}\right)=\beta_0+\beta_1 NCLB_s+\beta_2 Med_s+\beta_3 Wave_t+\beta_4 NCLB_s\times Wave_t + \beta_5 Med_s\times Wave_t+\lambda_1 StateVar_{s,t}+\lambda_2 IndVar_{i,s,t} \quad (1)$$

The variable and indices definitions for Eq (1) follow:

Indices: *i*- individual, *s*- state, *t*- time period (or NSCH wave)

ln: natural logarithm (base *e*)

$ADHD_{i,s,t}$: binary variable indicating whether child had ever been diagnosed with ADHD

$NCLB_s$: binary variable indicating state had No Child Left Behind-initiated consequential accountability

Med_s: binary variable indicating state had psychotropic medication law

$Wave_t$: binary variable indicating the later NSCH wave used in the analyses

$StateVar_{s,t}$: time varying state-level variables (i.e., number of healthcare providers per capita by age)

$IndVar_{i,s,t}$: individual-level variables (e.g., child's gender)

Eq (1) includes two difference-in-differences (DiD) terms: $NCLB_s \times Wave_t$ and $Med_s \times Wave_t$. The remainder of this appendix focuses on the latter DiD term, but the discussion also applies to the former.

The DiD term, $Med_s \times Wave_t$, estimates the association between changes in ADHD diagnostic prevalence and psychotropic medication laws. The intervention variable—Med_s—is time invariant, because the differenced model compares how ADHD diagnostic prevalence changed between two NSCH waves for two groups of states. The DiD parameter is β_5.

Eq (2) and Eq (3) show the parameters that estimate how the ln-odds of ADHD diagnostic prevalence changed between two NSCH waves in states that had a psychotropic medication law (Eq. 2) versus states without such law (Eq. 3). The term x includes all of terms in Eq (1), except for Med_s, $Wave_t$, and $Med_s \times Wave_t$.

$$\ln\left[\frac{p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=1,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=1,x}\right]- \ln\left[\frac{p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=0,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=0,x}\right]=\beta_3+\beta_5 \qquad (2)$$

$$\ln\left[\frac{p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=1,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=1,x}\right]- \ln\left[\frac{p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=0,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=0,x}\right]=\beta_3 \qquad (3)$$

When Eq (3) is subtracted from Eq (2) and each side of the equation is exponentiated (exp), then exp(B_3) is the result, which is the ratio of two odds ratios (see Eq. 4). In model 4, this parameter is estimated to be .75 (p=.04).

$$\frac{\dfrac{\dfrac{p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=1,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=1,x}}{\dfrac{p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=0,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=1,Wave_t=0,x}}}{\dfrac{\dfrac{p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=1,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=1,x}}{\dfrac{p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=0,x}{1-p(ADHD_{i,s,t}=1)\mid Med_s=0,Wave_t=0,x}}}=\exp(\beta_5) \qquad (4)$$

Odds ratios are difficult to interpret, let alone the ratio of two odds ratios. Therefore, we transform the DiD result in Eq (4) into probability units using the *margins* command in Stata 12 (1, 3). The result is a marginal effect that is estimated for each child in probability units, as shown in Eq (5). These marginal effects are averaged across the sample to obtain an average marginal effect. Importantly, the four terms in Eq (5) are the same as the four terms in Eq (2) and Eq (3), except the DiD is now in probability units instead of ln-odds units.

$$
\begin{aligned}
&Marginal\,Effect_{i,s,t} = \\
&([\hat{p}(ADHD_{i,s,t}=1)\,|\,Med_s=1, Wave_t=1, x] - [\hat{p}(ADHD_{i,s,t}=1)\,|\,Med_s=1, Wave_t=0, x]) - \\
&(\,[\hat{p}(ADHD_{i,s,t}=1)\,|\,Med_s=0, Wave_t=1, x] - [\hat{p}(ADHD_{i,s,t}=1)\,|\,Med_s=0, Wave_t=0, x])
\end{aligned}
\quad (5)
$$

In order to estimate the marginal effect in Eq (5), four predicted probabilities of having been diagnosed with ADHD are estimated for each child, by using the different values of the psychotropic medication law (Med_s) and NSCH wave ($Wave_t$) variables. For example, Eq (6) estimates the predicted probability for the first term in Eq (5), which sets Med_s = 1 and $Wave_t$ = 1. Therefore, Eq (6) sets Med_s = 1 and $Wave_t$ = 1 for every child in the sample, while each child's other covariate values retain their actual values. In Eq (6), the term $\hat{\delta}x$ includes all of terms in Eq (1), except for the following terms: $\beta_1 Med_s$, $\beta_2 Wave_t$, and $\beta_3 Med_s \times Wave_t$.

$$
\left[\hat{p}(ADHD_{i,s,t}=1)\,|\,Med_s=1, Wave_t=1, x_{i,s,t}\right] = \left[1+\exp\left(-\left(\hat{\beta}_2+\hat{\beta}_3+\hat{\beta}_5+\hat{\delta}x_{i,s,t}\right)\right)\right]^{-1} \quad (6)
$$

For each child, the remaining three predicted probabilities that he or she had been diagnosed with ADHD are calculated by setting Med_s = 1 and $Wave_t$ = 0; Med_s = 0 and $Wave_t$ = 1; and Med_s = 0 and $Wave_t$ = 0, respectively. The parameter estimates used to calculate the marginal effect for each child in Eq (5) are shown in Eq (7).

$$
\begin{aligned}
Marginal\ \ Effect_{i,s,t} = &\left(\left[1+\exp\left(-\left(\hat{\beta}_2+\hat{\beta}_3+\hat{\beta}_5+\hat{\delta}x_{i,s,t}\right)\right)\right]^{-1} - \left[1+\exp\left(-\left(\hat{\beta}_2+\hat{\delta}x_{i,s,t}\right)\right)\right]^{-1}\right) - \\
&\left(\left[1+\exp\left(-\left(\hat{\beta}_3+\hat{\delta}x_{i,s,t}\right)\right)\right]^{-1} - \left[1+\exp\left(-\left(\hat{\delta}x_{i,s,t}\right)\right)\right]^{-1}\right)
\end{aligned}
\quad (7)
$$

The average marginal effect is estimated using the weighted average of the marginal effects calculated for each child using Eq (7). The weights are sampling weights provided by NSCH. The standard error of the average marginal effect is estimated by Stata's *margins* command using the delta method (1), adjusted for clustering at state level, as was the case for estimating the logistic regression model in the main article.

II.B. Reporting the Average Marginal Effect of a Difference-in-Differences Result

As stated above, in order to estimate the marginal effect in Eq (5), four predicted probabilities of having been diagnosed with ADHD are estimated for each child, by using the different values of the psychotropic medication law (Med_s) and NSCH wave ($Wave_t$) variables. In Table 4 (main article), we report the weighted average of each of the four predicted (i.e., adjusted) probabilities, based on model 4 in the main article. The average marginal effect of the DiD term can be calculated by differencing these averages.[1]

From 2003-2011, children ages six to 13 residing in states with a psychotropic medication law had their adjusted prevalence decrease 4% (from 8.1% to 7.8%), but demographically similar children residing in states without a psychotropic medication law had their adjusted prevalence increase by 23% (from 8.1% to 10.1%).[2] This pattern results in a DiD average marginal effect of -2.2 percentage points (p=.02): (7.8% – 8.1%) – (10.1% – 8.1%). The p-values of the DiD average marginal effect (.02) and the DiD ratio of two odds ratios (.04) reported above are similar.

This DiD result is graphically displayed in Figure 1 (main article), which shows an adjusted ADHD diagnostic prevalence trend that is slightly decreasing in states with a psychotropic medication law, in contrast to an increasing trend in states without such law. Recall from the main article, these 2011 adjusted prevalences are understated, because

[1] When Stata calculates the average marginal effect of the DiD term, it calculates weighted average of child-level differences in Eq (7). This produces the same result as differencing the four weighted average probabilities reported in Table 3 (main article).
[2] Numbers presented in the text rounded, but calculations are based on more precise numbers.

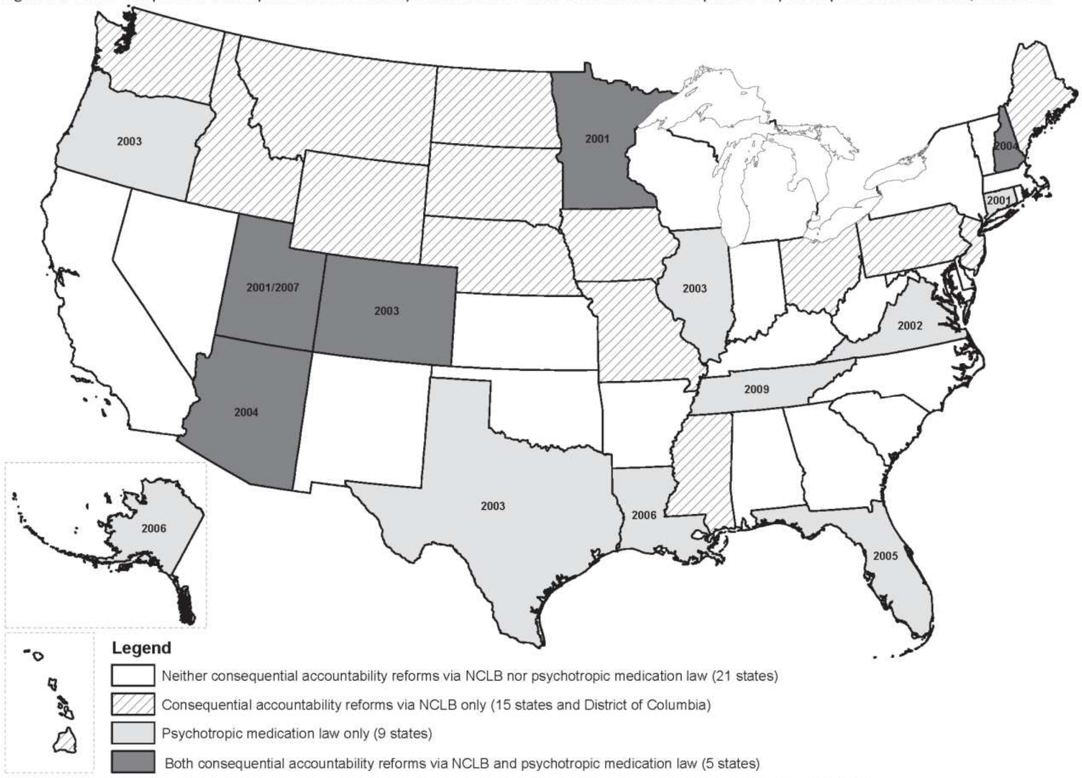
they excluded children diagnosed before the 2003 NSCH or before age 5; however, it is still valid to compare the adjusted prevalence changes from 2003 to 2011 between these two groups of states.

III. References

1. Williams R. Using the margins command to estimate and interpret adjusted predictions and marginal effects. Stata Journal. 2012;12(2):308-31.
2. Wooldridge JM. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press; 2002.
3. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp; 2011.

Online Supplement

Figure S1: States' Adoption of Consequential Accountability Reforms via No Child Left Behind and Adoption of Psychotropic Medication Laws, 2001-2012



Notes: The year listed in a state is the effective year of the psychotropic medication law. Utah has passed two laws, so two years are listed.
NCLB: No Child Left Behind Act of 2001