

Online Appendix Materials

Adequate (or Adipose?) Yearly Progress: Assessing the Effect of "No Child Left Behind" on Children's Obesity by Patricia M. Anderson, Kristin F. Butcher, and Diane Whitmore Schanzenbach

Appendix A: Data Construction

School-level Rates of Obesity and Overweight

An important part of Act 1220 of the 2003 Arkansas General Assembly was the weighing and measuring of all public school children. While the individual reports were confidential and only sent home to the student's parents, a report that includes school-level results is produced annually by the Arkansas Center for Health Improvement (ACHI). At a minimum, the percent of students who are overweight or obese, versus those who are normal or underweight, is reported.¹ For some schools, in some years, there is a complete breakdown into the four possible weight categories, while other years group overweight and obese together. Based on these reports, then, we create panel data on school-level rates of overweight or obese from 2004 to 2010.

School Academic Performance Reports

One of the requirements of No Child Left Behind (NCLB) is to make available school-level information on academic achievement in math and literacy. School report cards for Arkansas were obtained from the Department of Education for the academic years ending in 2002 through 2010. These school report cards provide information on the percent of students in

¹ Weight categories are based on age-by-gender percentiles from a fixed population, where underweight is below the 5th percentile, overweight is above the 85th percentile, and obese is above the 95th percentile.

each tested grade scored as proficient on the literacy test and the percent scored as proficient on the math test. These percentages are reported not only for the grade overall, but for the required subgroups as well. The groups we include for the analysis are whites, African Americans, Hispanics, and economically disadvantaged students. We exclude students with disabilities and with limited English proficiency. Subgroup proficiency rates are reported on the school report cards for groups with a sample size of at least 10, but only subgroups that have sample size of 40 (or 5 percent of total enrollment, whichever is larger) count toward the official accountability rating. Starting in 2008, the data we received contain a count of the number of students in a subgroup. For the prior years, we are able to estimate population sizes from the Common Core of Data (CCD). The CCD data report annual school-by-grade enrollment overall and for several of the subgroups of interest (whites, African Americans, and Hispanics). In addition, we were able to proxy for the number of economically disadvantaged students in each grade by multiplying the school-level fraction of students on free or reduced-priced lunch by the grade-specific enrollment. We then used group size from either the CCD or the report card data to omit tested groups from accountability calculations if they were based on too small of a population (and therefore did not count toward accountability status under Arkansas rules). No information is available from either source for the number of students with disabilities or limited English proficiency. After trying several approaches to proxy for whether the sample sizes in these two groups would have been large enough to count toward AYP, we dropped both of these subgroups from the analysis. Note that in the end this improved our predictive power when we modeled the official AYP status as coded by the state as a function of the program rules and the public-use data.

We next turn these school-subgroup-grade-test-year data into a school-year panel. For each school-grade-year, for each test, for each subgroup with a group size large enough to

count toward the official accountability rating, we standardize the proficiency rates around that year's AYP threshold (see table A.1). For example, for the fourth-grade math test, the 2006 threshold is 40 percent. If a subgroup had a 45 percent proficiency rate, their standardized rate for fourth grade math in 2006 is 5. Similarly, if a subgroup had a proficiency rate of 30, their standardized rate for fourth grade math in 2006 is -10. Thus, positive standardized rates represent meeting the AYP goal, while negative ones represent failure to meet AYP. Since achieving AYP requires every subgroup to meet the goal, we define a school-level minimum-group proficiency score as the minimum standardized rate across all grades and subgroups in the school for that year. We also maintain the minimum-across-grade math and literacy proficiency rates for the overall school population, as they reflect more generally on the school's overall academic performance. Additionally, based on the group-size information, and that includes school-level results, we calculate the percentage of the school's total students who are nonwhite, and who are economically disadvantaged, order to be able to control for observable demographics. The result is a school-year panel on AYP performance under NCLB and basic demographics.

The Final Analysis Sample

The obesity rate and AYP panels are merged by school and year. This gives us a full panel from 2004 through 2010. However, our main models will use not only lag of AYP performance (which is available for 2003), but also the lag of the obesity rate (which is not). Thus, our analysis sample effectively starts in 2005 and contains 4,588 school-year observations, representing 935 unique schools. Note the panel is unbalanced mainly due to some schools being too small in some years, although occasionally not all of the data were reported,

requiring the school-year to be dropped. Of the 935 schools, about two-thirds (613) enroll some students below grade 7.

The key variables for the analysis represent whether the school should feel pressured by NCLB to make marginal changes that may affect student weight. Our assumption is that there are some schools with proficiency rates well above the AYP goals, which should not feel pressured by NCLB at all. Similarly, there are some schools with proficiency rates well below the AYP goals, which may feel pressure but are unlikely to feel that marginal changes will relieve this pressure. Schools with proficiency rates close to the AYP goals, though, should be those most likely to feel pressured by NCLB to make marginal changes such as reducing recess time, having reward parties with unhealthy snacks, etc. Thus, we define pressure as having a scaled proficiency rate for the worst performing subgroup of -5 to 5. For our main specification, once a school has been coded as under pressure, it remains that way given that schools are unlikely to immediately undo small policy changes. For a school that has never been under pressure, current status is used to categorize the school as top or bottom. In three alternative specifications, we define pressure based on any subgroup (not just the lowest scoring) being within 5 points, and alternatively, we allow schools to alternate back and forth between being under pressure and not being under pressure (using both definitions).

Table A.1. Proficiency Rate Goals in Arkansas, by Grade Span and Year

	Grades K-5		Grades 6-8		Grades 9-12	
	Literacy	Math	Literacy	Math	Literacy	Math
2002	31.80	28.20	18.10	15.30	19.50	10.40
2003	37.48	34.18	24.93	22.36	26.21	17.87
2004	43.16	40.16	31.76	29.42	32.92	25.34
2005	48.84	46.14	38.59	36.48	39.53	32.81
<i>2006 amendment changed schedule</i>						
2006	42.40	40.00	35.20	29.10	35.50	29.20
2007	49.60	47.50	43.30	38.06	43.56	38.05
2008	56.80	55.00	51.40	47.02	51.62	46.90
2009	64.00	62.50	59.50	55.98	59.68	55.75
2010	71.20	70.00	67.60	64.55	67.75	64.60
2011	78.40	77.50	75.70	73.41	75.81	73.45
2012	85.60	85.00	83.80	82.28	83.88	82.30
2013	92.80	92.50	91.90	91.14	91.94	91.15
2014	100	100	100	100	100	100

Table A.2. Summary Statistics by Pressured Status

	Full Sample (1)	Pressured in Past (2)	Never Pressured (3)	Above Margin (4)	Below Margin (5)
Pressured in Past	0.553 (0.497)	1	0	0	0
Below Pressured Margin	0.239 (0.427)	0	0.535 (0.499)	0	1
Above Pressured Margin	0.207 (0.406)	0	0.465 (0.499)	1	0
Overweight Rate	38.44 (6.597)	38.92 (6.281)	37.85 (6.925)	34.39 (6.475)	40.85 (5.819)
Percent Nonwhite	29.99 (28.74)	24.45 (25.24)	36.85 (31.24)	15.26 (15.99)	55.59 (29.07)
Percent Economically Disadvantaged	57.15 (19.78)	56.48 (17.57)	57.98 (22.20)	46.04 (17.52)	68.35 (20.57)
English Proficiency Rate Relative to AYP Threshold (Previous Year)	9.319 (17.27)	9.513 (15.10)	9.079 (19.63)	24.31 (10.72)	-4.136 (15.63)
Math Proficiency Rate Relative to AYP Threshold (Previous Year)	11.48 (18.14)	12.38 (14.70)	10.36 (21.60)	28.38 (10.83)	-5.270 (15.61)
Overweight Rate (Previous Year)	38.27 (6.264)	38.69 (5.979)	37.76 (6.565)	34.30 (6.031)	40.75 (5.443)
Observations	4,588	2,539	2,049	952	1,097

Notes: Pressured in Past means that the minimum-scoring subgroup had a proficiency rate within 5 points of the AYP target for some year in the past. Above AYP margin implies that the minimum-scoring subgroup has never had a proficiency rate within 5 points of the AYP target, and had a proficiency rate more than 5 points above the AYP target last year. Below AYP margin implies that the minimum-scoring subgroup has never had a proficiency rate within 5 points of the AYP target, and had a proficiency rate more than 5 points below the AYP target last year. Overweight Rate includes all weights above normal weight. Standard deviations in parentheses.

Table A.3. Comparing Lagged Dependent Variable to Fixed Effects Results

Pressure Definition:	Minimum Subgroup, Any Past Pressure		Any Subgroup, Any Past Pressure	
	Lagged dependent variables	Fixed effects	Lagged dependent variables	Fixed effects
	(1)	(2)	(3)	(4)
Pressure Indicator	0.522*** (0.151)	0.436** (0.222)	1.129*** (0.195)	0.472* (0.246)
Sample Size	4588	6625	4588	6625
R-squared	0.496	0.001	0.498	0.001

Notes: Columns 1 and 2 differ from columns 3 and 4 in how NCLB pressure is defined. Minimum Subgroup and Any Past Pressure (columns 1 and 2) define the pressured group by whether a school's lowest scoring subgroup had a proficiency rate within 5 points of the AYP target for some year in the past. Any Subgroup and Any Past Pressure (columns 3 and 4) are indicators for whether any accountability subgroup in the school had a proficiency rate within 5 points of the AYP target in any prior year. Columns 1 and 3 differ from columns 2 and 4 in specification approach. The odd-column models are identical to those in table 4, and include the lagged rate of overweight, plus a quartic in percent nonwhite, a quartic in percent economically disadvantaged, a quartic in the standardized overall literature proficiency rate, a quartic in the standardized overall math proficiency rate, and an annual trend. The even-column models include school-level fixed effects. Standard errors which are robust to heteroskedasticity and within-school correlation are in parentheses.

*p<0.1; **p<0.05; ***p<0.01.

Appendix B: Determining Adequate Yearly Progress in Arkansas

During the time period studied, high-stakes math and literacy tests were initially given in grades 4, 6, and 8, and in algebra, geometry, and eleventh grade literacy. Math and literacy tests were gradually expanded to include grades 5 and 7. The publicly available school report card data report the percentage of students with scores on each test in various ranges (below basic, basic, proficient and advanced), separately for math and literacy. For each school, this information is reported not only for the “combined population,” but also for a variety of subgroups defined by race/ethnicity (African-American, Caucasian, Hispanic) and special classification status (economically disadvantaged, limited English proficiency, students with disabilities). Average proficiency rates are reported in the public-use data for all group-by-test cells with a sample size of at least 10. However, the cell does not count toward proficiency unless there are at least 40 students in the cell (for schools with fewer than 800 students) or the cell represents at least 5 percent of enrollment (for schools with 800 or more students).

If fewer than 95 percent of students are tested in math and reading, the school automatically fails AYP. The report card lists whether the school meets the 95 percent threshold, and if not it reports the school-wide percentage of students tested. The denominator of this calculation is all students enrolled in the grade or course at the time of testing. If the proficiency rate in the cell is greater than or equal to that year’s passing target (termed “Annual Measurable Objective” or AMO and displayed in table A.1), and at least 95 percent of the eligible students are tested, then that group-by-test meets AYP. The school as a whole meets AYP if both its combined population and all subgroups that count toward proficiency meet the AYP requirements. If any subgroup fails to meet that year’s AMO (and does not meet AYP through the safe harbor provision described below), then the school has failed to meet AYP. As a result, the binding constraint is the proficiency rate of the worst-

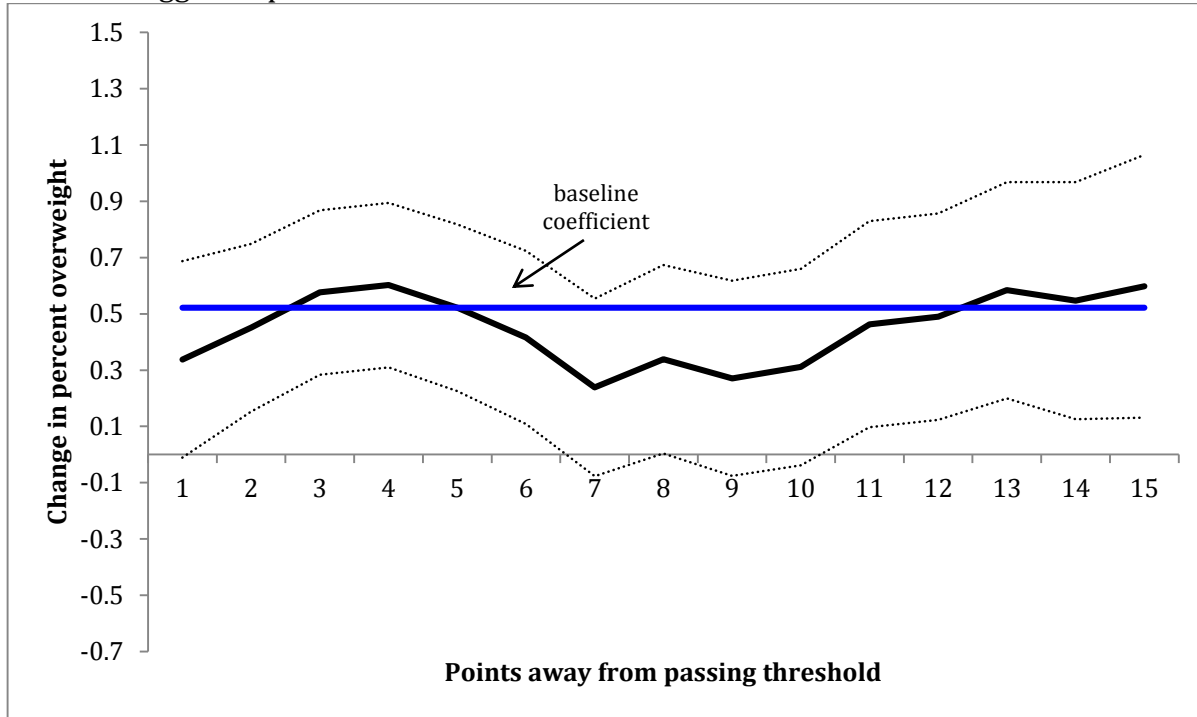
performing subgroup that is large enough to count toward the rating. Our empirical approach reflects this by defining each school by the performance of its subgroup with the lowest passing rate (see figure B.1).

A school can have its AYP status determined based either on the current year's performance or the average over the prior three years. The metric chosen can vary across years, but must be consistent across all subgroups within the same year. Schools can also be classified as meeting AYP requirements even if the passing rate is below the year's AMO under the safe harbor provision, as long as at least 95 percent of students are tested in the year. Under safe harbor, if the fraction of students scoring below proficient in each subgroup declines by at least 10 percent from the prior year, then the school meets AYP requirements for that year. In addition, the school meets AYP under safe harbor if its observed increase in proficiency rates falls within a 95 percent confidence interval of the safe harbor goal. Over this time period, the minimum increase necessary to meet AYP under safe harbor was always greater than 5 points, which is the cutoff for being a "pressured" school in the analysis.

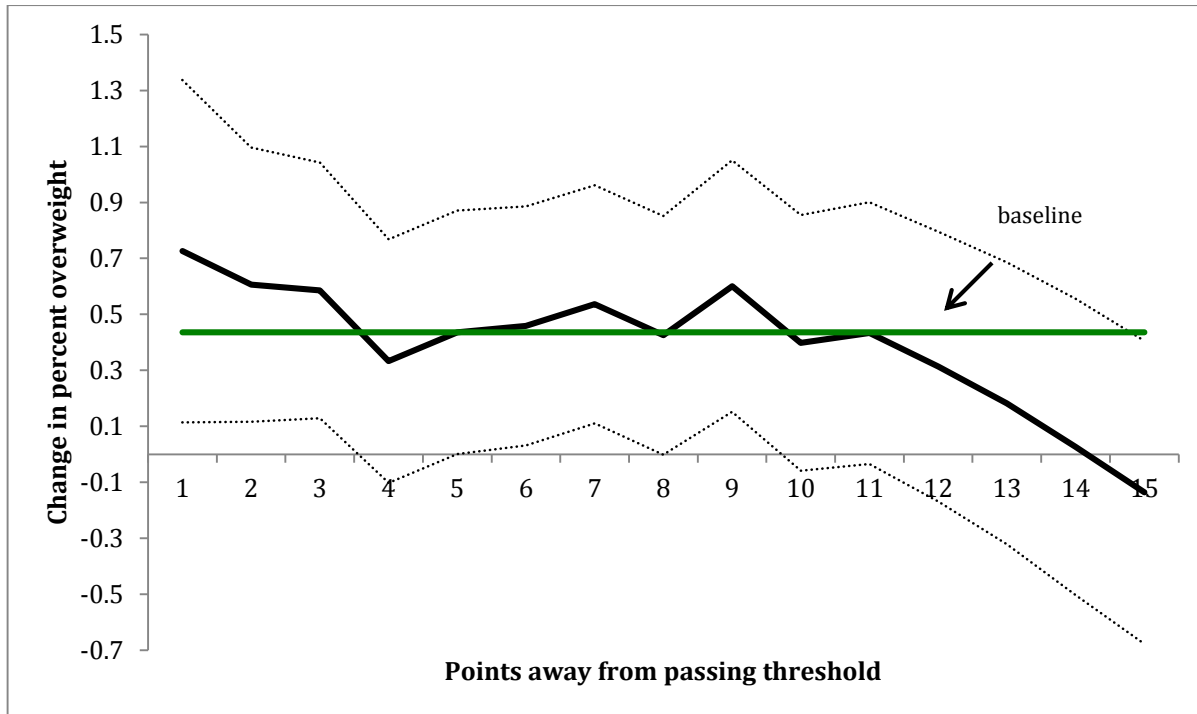
In order to meet AYP, schools must also satisfy the appropriate "secondary indicator." For high schools, graduation rate must be at least 70 percent, and for non-high schools attendance rates must be at least 91.13 percent. Among schools that are passing due to test rates, relatively few fail due to the secondary indicator. We do not use this information in our analysis.

Figure B.1. Estimated Effect Sizes When Definition of “Close to Threshold” Varies

Panel A: Lagged Dependent Variable Models



Panel B: Fixed Effects Models



Notes: The top panel shows the estimated coefficient (and 95% confidence interval) on the pressured indicator in an equation analogous to equation 1, where the definition of the pressured indicator ranges from 1 point above or below the NCLB threshold to 15 points above or below it. The baseline case is 5 points above or below the threshold, and is also shown in table 2, column 4. The bottom panel shows analogous results using school-level fixed-effects models.