Preliminary Version
Comments Welcome

The Effect of School Accountability Policies on Children's Health

Patricia M. Anderson
Dartmouth College and NBER

Kristin F. Butcher
Wellesley College

Diane Whitmore Schanzenbach
University of Chicago

October 2009

Note: We thank Jannine Riggs, Denise Airola, and Jim Boardman of the Arkansas Department of Education, RWJ grant number #57922 and the Rockefeller Center at Dartmouth College Reiss Family Faculty Research Grant for financial support, and Brian Dunne, A.J. Felkey, Brenna Jenny and Alan Kwan for excellent research assistance.

I. Introduction

Childhood obesity has increased dramatically over the past three decades, from about 4 percent during the mid-1970s to 16 percent today. There have been many changes in children's lives during the period when children's obesity has been increasing (see Anderson and Butcher 2006a). In particular, there have been changes at home and at school that may contribute to increased obesity. Understanding how the school environment may contribute to obesity is critical as the school environment may be relatively more within the control of public policy makers than the family. In recent years, especially, pressures on schools have changed dramatically. First, over the 1980s, many states passed tax and expenditure limitation laws that fundamentally altered the way schools had traditionally been financed. Second, during the 1990s and 2000s, the emphasis has turned to "accountability," which puts new pressure on schools to improve academic outcomes, without necessarily providing more resources with which to produce these outcomes. In 2002, the Federal No Child Left Behind (NCLB) legislation was passed, requiring states to define and implement stringent accountability standards and prescribing increasing penalties for schools that fail to meet their state's standard.

This paper investigates how these accountability pressures may affect children's obesity. Children's health is typically not among the outcomes for which schools are held accountable – standard test achievement is the primary area monitored, with secondary emphasis on attendance and graduation rates. Schools facing increased pressures to produce academic outcomes may reallocate their efforts in ways that have unintended consequences for children's health. The new financial pressures due to accountability rules may, for example, induce school administrators to try to raise new funds through outside food and beverage contracts,[1] or time pressures may cause them to cut back on

[1]Anderson and Butcher (2006b) find evidence that schools that are under more financial pressure are more likely to give students access to junk food and that students in these schools have higher BMI.

recess and physical education in favor of increased academics.[2] Additionally, there is the possibility that testing pressures increase cortisol secretions in children due to the increased stress, which can impact obesity.

We examine the impact of accountability programs in two states. First, we create a unique data set for Arkansas that allows us to test the impact of NCLB rules on students' weight outcomes. These data combine school-level rates of "obesity" and "overweight" for children in all schools in Arkansas with data from the Arkansas Department of Education on standardized test proficiency rates in English and math for all schools, by grade and subgroup.[3] The standardized test pass rates are those used for determining whether a school is making adequate yearly progress (AYP) under NCLB. We create a similar data set for California, using two sets of data from the California Department of Education. School-level information from the body composition area of the physical fitness test administered to all 5th, 7th and 9th graders in the state is matched to school-level data on proficiency in English and math. Because California has a state accountability system based on the academic performance index (API) that goes beyond the AYP requirements of NCLB, we also include basic information on the API. Finally, we survey school principals in Arkansas about their schools' responses to the implementation of NCLB.

Below, we describe both the Arkansas and California data in more detail, and discuss our approach to modeling the role that accountability pressures may play. We then present results from empirical models on obesity rates for each state, as well as some preliminary survey results, before concluding.

[2] Center on Education Policy (2007) finds 20% of school districts have decreased recess time since NCLB was enacted, with an average decrease of 50 minutes per week.

[3] Obesity is defined as having a body mass index (BMI) greater than the 95th percentile of a distribution of age- and sex-specific BMIs from a baseline population from the 1970s. Overweight is defined analogously, with BMI greater than the 85th percentile. The official Arkansas documentation follows CDC convention and labels these thresholds differently as "overweight" and "at risk of overweight," respectively. We will use the more common terms "obese" and "overweight" instead.

II. Background and Data

A. Arkansas Assessment of Childhood and Adolescent Obesity[4]

In 2003 the state of Arkansas passed a sweeping act intended to help combat childhood and adolescent obesity. Although obesity has been increasing nationwide, obesity levels were particularly high in Arkansas. In 2003, about 21 percent of school aged children in Arkansas were obese or overweight, while this figure was about 18 percent for the nation as a whole.[5] A multifaceted coalition came together to address the challenge of childhood obesity, and passed Act 1220 of the 2003 Arkansas General Assembly.[6] Reporting health risk information, in particular a child's BMI and whether that BMI indicated the child was underweight, normal weight, overweight, or obese, to each parent was a central component of this initiative (ACHI 2004).

The Arkansas Center for Health Improvement spearheaded the effort to collect height and weight information for each school child in the state of Arkansas. This effort included ensuring that each school had the equipment and trained personnel necessary to accurately weigh and measure each child.[7] A letter then went home to each parent describing the child's BMI, where this fit in the BMI distribution, and the type of health risks that might be associated with the child's BMI. Parents with children with an unhealthy weight were urged to consult a physician. An implicit assumption of this effort was that if better information was in the hands of parents, they could make, or help their children make, better informed, more healthful, choices that would improve their weight outcomes. Additionally, an annual public report is produced (available on the Arkansas Center for Health Improvement website) with the percent of students who are underweight, normal weight, overweight,

[4] This section draws heavily from the yearly reports on the Arkansas Assessment of Childhood and Adolescent Obesity released by the Arkansas Center for Health Improvement. Reports are available online at: www.achi.net

[5] Comparison of Table 1 in ACHI (2004) to NHANES 2003-2004 calculations (http://www.cdc.gov/nccdphp/dnpa/obesity/childhood/prevalence.htm).

[6] The coalition included parents, school nurses, teachers, and administrators, private foundations (including Robert Wood Johnson), physicians, hospitals, universities, Governor Mike Huckabee, the Arkansas Departments of Education and Health, among many others.

[7] Training included taking each measure a number of times to ensure accuracy.

and obese at each public school in Arkansas. Thus, due to the Arkansas Assessment of Childhood and Adolescent Obesity, we have panel data on school-level obesity rates from 2004 to 2007.

B. California Physical Fitness Test

Statewide physical fitness testing in California is part of the 1995 California Assessment of Academic Achievement Act.[8] Results from the physical fitness test (PFT) are reported on the annual School Accountability Report Card (SARC). Thus, in California, the accountability system does take into account the students' health. That said, while there is reporting of the PFT results, there do not appear to be any real repercussions for a school based on the number of students being found to not be in the healthy fitness zone (HFZ). Thus, the possibility remains that the academic accountability rules may have spillovers on student health. A battery of tests developed by The Cooper Institute called *FITNESSGRAM*© is used for the California PFT. There are six fitness areas: Aerobic Capacity, Abdominal Strength and Endurance, Body Composition, Trunk Lift and Flexibility. We focus only on the Body Composition area. While the Cooper Institute recommends using skin-fold measurements for determining whether body composition is in the healthy zone, they allow the use of BMI, and as best as we can tell, that is the typical measure used by the California PFT. Unlike in Arkansas, where the percent overweight/obese are based on the CDC $85^{th}/95^{th}$ percentile cutoffs, the Cooper Institute provides software to schools with different cut points that are meant to approximate having over 25 percent body fat for boys or 32 percent for girls (Welk and Meredith, 2008).

School-level results from the PFT are available from the California Department of Education website. Recall that the only tested grades are 5^{th}, 7^{th} and 9^{th}. Most schools are configured such that only one of these grades is present, and we use this tested grade as representative of the entire school. For schools in which more than one grade is present, for simplicity we just use data from the youngest grade. We then assume that the fraction of students not in the healthy zone for the body composition

[8] Testing was actually first authorized in 1976, but was reestablished in 1995 (see California Department of Education website at http://www.cde.ca.gov/ta/tg/pf/pftprogram.asp)

test can be considered to be the fraction of students who are overweight or obese. Recall that this definition will not exactly match that used in Arkansas.

C. School Academic Performance Reports

While some states (such as California) began implementing academic accountability earlier, the Federal government began its nationwide effort to improve children's academic outcomes in 2002, through the NCLB school accountability policies. Although the details of NCLB as it is implemented are left up to each state, and can be quite complicated, the basics of the policy in both Arkansas and California can be described fairly straightforwardly. Students in selected grades take state-wide standardized tests in reading and math that are appropriate to their grade. The state sets the passing score (e.g. a score above which a student is determined to be "proficient") on the test and creates a schedule showing what fraction of students must pass each test in each year for the school to be considered to have made Adequate Yearly Progress (AYP). The passing rate increases over time until 100 percent of students are required to meet the standard by 2014.

A school's AYP designation is determined by the average passing rate of its students overall. In addition, the passing rate of all designated sub-groups that have a large enough must also meet the goal.[9] Student sub-groups are defined by race (for whites, African Americans, Hispanics, etc.), and as low socio-economic status, English language learners, migrants, and students with disabilities. If any one of the student subgroups fails to attain AYP, then the entire school is designated as failing to meet AYP.[10]

While these basic rules are straightforward enough, in practice a school can be deemed to meet or fail to meet AYP for several other reasons. For example, even if a school (or subgroup) has a lower fraction of students meeting AYP than the passing standard requires, it still might make AYP through

[9] The definition of "large enough" varies by state. In Arkansas, there must be 40 students. In California, there must be either 50 students making up 15 percent of the total, or at least 100 students.

[10] We refer to these as "failing" schools, though the official nomenclature is that these schools are in "School Improvement Status."

the "Safe Harbor" provision, which allows a school to be deemed as passing if the percentage of failing students (within subject and subgroup) declines by ten percent relative to the prior year. On the other hand, a school will be deemed as failing despite its passing rate if too low a fraction of its students participate in the test, or if attendance or graduation rates are below the target threshold.[11] Because of the many details involved in determining AYP status, in order to perfectly predict a school's status it is necessary to use micro-data on student level performance. Since only aggregate data are available to us, we cannot perfectly predict AYP status. Nonetheless, the aggregate data should be sufficient to identify schools that should think that they are just marginally making or missing AYP.

In California, running in parallel to the federal requirements for meeting AYP are the state requirements on API. The API is an index between 200 and 1000 that reflects performance on a set of statewide tests. For students not yet in high school, over 50% of this composite score comes from the English test and another third from math, with a small fraction coming from science for elementary school students and also history for middle school students. For high school students, the composite is somewhat more evenly weighted across those fields, although there is an additional set of English and math tests that count. Given a school's current API (base API), a target is set for the coming year (growth API). For a base API between 200 and 690, the school's growth target is .05*(800 – base API). For a base API between 691 and 795, the growth target is 5 points. For each additional base API between 796 and 799, the growth target falls by 1 point. Finally, schools with a base API of 800 must maintain that API. As with AYP, this calculation is done for both the school overall, and for large enough subgroups.

School report cards for both states are provided by the Department of Education. These school report cards provide information on the percent of students scored proficient on the English test and the percent proficient on the math test. These proficiency rates are provided for both the school overall,

[11] More information on the Arkansas accountability plan is available at http://arkansased.org/nclb/pdf/accountability_wkbk_021208b.pdf.

and for the school's subgroups. For California, the school's base API and growth API are also provided. Because subgroup proficiency rates are reported on the school report cards for groups that are smaller than those used for accountability, we need to obtain measures of subgroup sample sizes. Although perfect data on this are not publicly available, we are able to estimate population from the Common Core of Data (CCD) for the years 2002-2006.[12] The CCD data report annual school-by-grade enrollment overall and for several subgroups of interest (whites, African Americans and Hispanics). In addition, we were able to proxy for the number of low-income students in each grade by multiplying the school-level fraction of students on free or reduced-priced lunch by the grade-specific enrollment. We used the information on enrollment from the CCD to omit test scores from accountability calculations if they were based on too small of a population.[13]

III. Methodology

We take two basic approaches to investigating the possibility that accountability pressures may be contributing to childhood overweight. One approach assumes that accountability-induced behavioral changes are likely to be greatest among schools that are close to meeting their targets. That is, schools easily meeting standards are unlikely to feel the need to change their behaviors in the face of accountability, while schools very far from making the standards may feel pressure, but will be less likely to think that a small change such as a reduction in recess time will be useful in addressing their deficiencies. However, we expect that schools with test scores just above and just below the target in year t-1 are the most likely to make the types of changes that might result in more overweight and obese students by year t. By comparing these "close" schools to those far away from the thresholds, we can determine if accountability is having an unintended impact on children's health. The other

[12] Because the BMI and test score data extend to 2007 but CCD data are not yet available for that year, we assign 2007 CCD data to be the same as the 2006 data.
[13] We also use the CCD data to create school-level demographics on percent nonwhite and percent poor (i.e. receiving free or reduced-price lunch).

approach assumes that if a school has seen larger proficiency gains over a year, then that school must have undertaken some types of behavioral changes to achieve those gains.

We start with Arkansas, by considering the AYP thresholds spelled out under NCLB, and determining what schools are "close" to meeting the standard. Since a school will fail to make AYP if any subgroup fails to meet the proficiency goal, we are especially interested in the worst-performing subgroup. For each school-grade year, for each test, for each subgroup with an acceptable group size, we standardize the proficiency rates around the AYP threshold. For example, for the 4th grade math test, the initial threshold is 40 percent. If a subgroup had a 45 percent proficiency rate, their standardized rate for 4th grade math in the initial year is 5. Similarly if a subgroup had a proficiency rate of 30, their standardized rate for 4th grade math in the initial year is -10. Thus, positive standardized rates represent meeting AYP, while negative ones represent failure to meet AYP. Since AYP is determined at the school-level, not grade-level, we then aggregate the data to the school-year level. We use the worst performing grade overall and for each subgroup to be representative of the school. We then choose the worst performing subgroup upon which to base our assessment of AYP performance. We also maintain the overall math and literacy rates, as they reflect more generally on the school's academic performance. Additionally, based on the CCD, we calculate the percentage of the school's students who are nonwhite, and the percent poor to control for observable demographics.

Our first approach, then, is to consider schools to be "marginal" if they have a minimum subgroup passing rate that is close to the AYP threshold. We define close as being 5 percentage points above or below the threshold.[14] While schools may have some idea that they are going to be close to making or missing the AYP threshold and change behaviors contemporaneously, we will nonetheless estimate current rates of overweight or obese based on the previous year's test results to ensure that the school has had time to react to being close to the AYP threshold. Our second approach focuses entirely on the overall school performance, not the subgroups. The school's change in math

[14] We have also experimented with alternative divisions with qualitatively similar results.

proficiency is added to the school's change in literacy proficiency to define the total proficiency gain from one year to the next. As with the model focusing on the marginal schools, we also control flexibly for current demographics and academic performance.

The approach using the data from California is similar, except we focus more on the school's progress toward API, since it appears the state accountability system is more binding that that imposed by NCLB.[15] To define a marginal school in California, we simply choose the 42 percent of schools that had an API growth target of exactly 5 points.[16] To focus on schools that achieved larger gains, we calculate the previous year's growth in the API, while controlling for what the API target is (and its square). As with Arkansas, all models control flexibly for basic demographics and overall school proficiency in math and English.

IV. Results

We start by examining some basic descriptive statistics on the final data sets. Looking at the top panel of Table 1, in the first column we see that almost a fifth of schools in Arkansas are classified as being "close" to the AYP threshold, with the remaining schools being almost evenly split between being more clearly in failing territory and passing territory. Overall, at the average school 38.3 percent of students are overweight or obese, 53.2 percent receive free or reduced-price lunch and 25.4 percent are nonwhite. Finally, this average school met the target English proficiency rate by 14.03 percentage points and the target math proficiency rate by 12.93 points, but over the last year the total proficiency rate actually dropped by over one half of a percentage point. The second and third columns, present characteristics for the average school which is "close" to meeting the AYP goal and the average school

[15] We initially planned to focus entirely on NCLB, so the data we collected from the California DOE only has the subgroup results for AYP and the overall API. We plan to obtain the full API data to more appropriately analyze the CA data in the near future.

[16] Note that if we use the AYP proficiency targets to define marginal schools as being within 5 points of AYP, there are 31 percent of schools within this range, and a full 55 percent meeting the AYP by more than 5 points. By contrast, 28 percent of schools need less than 5 points of API growth to meet their target and 30 percent need more than 5 points of API growth to meet their target.

which is not. While there is a slightly higher rate of overweight and obesity at the close schools, this does not appear to be simply due to demographics. These schools have a slightly lower rate of free and reduced-price lunch recipients and a much lower nonwhite percentage. In terms of overall school performance, the two types of schools are fairly similar - the overall test performance of these close schools is just a tiny bit worse than the non-close schools in Math, and is a bit better in English . Note however, the marginal schools did see more proficiency growth over the past year, up almost 3 percentage points, while the other schools lost almost 1.4 percentage points.

Turning to the second panel, we see that 42 percent of schools are labeled as marginal, with an API gain target of exactly 5 points, while of the inframarginal schools, 30 percent have a bigger target and 28 percent have a smaller target. Using the CA definition, just 32 percent of students at the average school are overweight or obese, while 54 percent receive free or reduced-price lunch and 68 percent are nonwhite. The average school met the AYP English proficiency rate by 20.6 percentage points and the math proficiency rate by 23.9 percentage points. In terms of API, the average school's target was simply to not lose more than 13 points, while the actual previous gain was 10.4 points. Again comparing the second and third panels, we see that the average marginal school has a slightly higher rate of overweight and obesity. Unlike in Arkansas however, this raw difference may be attributable to demographics, as there is a higher percentage of free and reduced-price lunch students and a slightly higher percentage of nonwhite students. Similarly, their overall proficiency rates in English and math are lower and they have a positive API growth target, but had slightly smaller gains over the last year. Recall, however, that in both states we will be flexibly controlling for demographics and overall proficiency rates, and the California API growth regression will also control for the API target.[17]

[17] Note that for AYP, all schools face the same target in a given year, so current proficiency is a sufficient to capture the task they face.

Table 2 presents the results of the basic models described above, with columns 1 and 2 being Arkansas, and columns 3 and 4 being California. Starting with column 1, we see that marginal schools in Arkansas have a rate of overweight that is just under 1 percentage point higher than the inframarginal schools. While not shown in the table, much of this effect is driven by a comparison with the schools having a proficiency rate that is more than 5 percentage points away from the AYP goal, as only the comparison with this group is significantly different from zero. The alternate approach, shown in column 2, also estimates a significant effect of NCLB. In this case, schools making a bigger proficiency gain over the past year are slightly more likely to have overweight and obese students. Because Arkansas set up its schedule for AYP to involve equal increments each year until reaching 100 percent proficiency, a typical school that just met their goal in one year will be expected to add about another 7 or 8 percentage points to each of the English and math proficiency rates, for a total gain of about 15. A school making this size gain, would thus be expected to have a rate of overweight/obesity that is about 0.4 percentage points higher than a school not making any gains.

In column 3, we look at the schools we have declared marginal in California. Here, the estimate is actually negative, albeit not significantly different from zero, so that we cannot rule out small positive effects.[18] With this API-based measure of marginal, however, we can rule out effects of the size seen with the AYP-based measure used in Arkansas. It remains to be seen if the same can be said about a model based on API that more closely parallels the AYP model from Arkansas. Recall that currently the California data do not allow us to properly check for subgroup level API goals, in the way that the worst-performing subgroup is used in the AYP models. Finally, in column 4, we estimate a model based on API-growth that is a bit more similar in spirit to the Arkansas model of proficiency

[18] While not shown in the table, in this case, this zero effect is due to offsetting effects. The schools with an API target of 5 have significantly higher rates of overweight and obesity than those with an API target below 5, and a significantly lower rate than those with an API target above 5. Note this result is not what would be expected simply from poor demographic controls, since lower target schools should be "better" in most dimensions. It may, however, imply that only those with very small API targets do not feel accountability pressures.

gains, since both are based on overall school performance, not subgroup performance. Here, we estimate a positive and significant effect, albeit one that is quite small. The average school in our sample manages about a 10 point gain in their API index from year to year, which would imply just a 0.1 percentage point increase in overweight students.

Overall, then, based on Table 2 there are some intriguing results hinting at the possibility that accountability pressures may have unintended side effects on children's health, but one cannot really point to a stable causal relationship. If one were to see clear impacts of our measures of accountability pressures on school behaviors, however, the current results might seem more reasonable. Thus, to investigate behavior changes at schools due to NCLB, we are currently surveying principals of our sample schools in Arkansas. As we currently only have about a 10 percent return rate, though, it is difficult to draw many conclusions. There is some evidence that schools with bigger changes in proficiency had larger reductions in minutes of recess or free time per week. When substituting recess/free time minutes for the obesity rate in the model from column 2 in Table 2, the coefficient on the change in the proficiency rate is -0.82 (implying about 1 minute less), and is significant at the 10 percent level. Granted, this effect does not seem especially economically significant, though. No other food (e.g. use of food rewards, increases in vending or beverage contracts, etc.) or activity variables are significant in the change in proficiency models, and no behaviors are ever significantly related to a school being close to the AYP threshold. As we continue to collect responses, we hope to be able to say something more precise about school behaviors in response to accountability.

V. Conclusions and Further Avenues for Research

Through the No Child Left Behind Act, schools face increasing pressure to deliver on standardized tests. Since schools are held accountable for test scores, but not for other student outcomes, such as children's health, schools facing pressure may make decisions designed to increase test scores that may have unintended consequences for children's weight. Previous research by Figlio

and Winicki (2005) has found that schools serve higher calorie meals on the days of high-stakes exams, for example. More broadly, schools trying to increase test passing rates may change the sort of foods available because food sales may generate profits that can be used on supplemental instruction. Additionally, there is evidence that since the passage of NCLB, schools have reallocated time toward math and literacy and away from other subjects. Most importantly, in allocating more time to instruction, they are reducing time for physical activity, like recess and physical education classes. These changes may inadvertently increase overweight and obesity.

Because it is clear that schools with lower test scores also have students with worse socioeconomic outcomes and thus tend to have worse weight outcomes, one cannot simply comparing poor performing schools to top performers. Instead, we take two approaches, both of which control flexibly for demographics and overall school academic performance. In the first approach, we compare schools who are likely marginal, in the sense that they expect to come very close to failing to meet accountability, with those that are inframarginal. These inframarginal schools expect either to easily make the standard or to have very little chance of doing so. Thus, it is the marginal schools that would be most likely to try changing their behavior in these small ways that might affect both their students' test scores and weight outcomes. In fact, we find that schools in Arkansas who were within 5 points of the AYP threshold (in either direction) in year t-1 have a relatively small, but statistically significantly higher rate of overweight and obesity in year t. Estimates for schools in California that have a 5 point goal increase for their API score are not significant. Our second approach assumes that schools that have made gains are the types of schools that might have made potentially deleterious behavioral changes. In both Arkansas and California, schools making bigger gains on the accountability standards have slightly higher percentages of overweight and obese students.

These results present prima facie evidence that the NCLB accountability rules may have unintended adverse consequences for student health. While these results are perhaps too preliminary to indict accountability in the childhood obesity epidemic, they do suggest that parents and school

administrators should keep in mind the potential for impacts on children's health as they consider how to reallocate school resources in pursuit of test score gains.

References

Anderson, Patricia M. and Kristin F. Butcher, "Trends in and Potential Causes of Childhood Obesity," *The Future of Children: Child Overweight and Obesity*, vol. 16, no. 1, Spring 2006a, Brookings Institution Press.

Anderson, Patricia M. and Kristin F. Butcher, "Reading, Writing, and Refreshments: Do School Finances Contribute to Childhood Obesity?," *Journal of Human Resources*, vol. 41, no. 3, Summer 2006b, pp.467-494.

Arkansas Center for Health Improvement, "The Arkansas Assessment of Childhood and Adolescent Obesity," 2004.

Arkansas Center for Health Improvement, "The 2005 Arkansas Assessment of Childhood and Adolescent Obesity: Online State Report," September 2005.

Arkansas Center for Health Improvement, "The Arkansas Assessment of Childhood and Adolescent Obesity – Tracking Progress: Online State Report, Year 3 (Fall 2005- Spring 2006)," 2006.

Arkansas Center for Health Improvement, "Assessment of Childhood and Adolescent Obesity in Arkansas: Year Four (Fall 2006 – Spring 2007)," 2007.

Center on Education Policy. *Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era*. Washington DC: 2007.

Figlio, David and Joshua Winicki, "Food For Thought: The Effects of School Accountability Plans on School Nutrition," *Journal of Public Economics*, Volume 89, Issues 2-3, February 2005, Pages 381-394.

Going, Scott B., Timothy G. Lohman, and Harold B. Falls, "Body Composition Assessment" in Fitnessgram/Activitygram Reference Guide (Gregory J. Welk. and Marilu D. Meredith eds.), Dallas, TX: The Cooper Institute.

Schanzenbach, Diane Whitmore "Does the Federal School Lunch Program Contribute to Childhood Obesity?," *Journal of Human Resources*, vol. 44, no. 3, Summer 2009, pp.684-709.

Table 1: Summary Statistics

Arkansas	Full Sample	Marginal Schools (w/in 5 pts of AYP)	Inframarginal Schools
W/in 5 points of AYP	.1762		
	(.3811)		
< 5 points from AYP	.3931		
	(.4885)		
> 5 points from AYP	.4307		
	(.4953)		
Overweight/Obese	38.2691	38.8424	38.1464
	(5.7548)	(5.0057)	(5.8965)
Free/Reduced-Price Lunch	53.2017	51.9177	53.4763
	(19.0556)	(16.6473)	(19.5248)
Nonwhite	25.3864	18.7605	26.8038
	(28.3841)	(24.3652)	(28.981)
English	14.0348	15.575	13.7053
	(16.6858)	(14.8394)	(17.0399)
Math	12.9291	12.7197	12.974
	(18.7973)	(14.896)	(19.5339)
Proficiency Gain	-.631	2.7252	-1.3542
	(26.7986)	(24.2245)	(27.2727)
Observations	2633	464	2169
California	Full Sample	Marginal Schools (5-Point API Goal)	Inframarginal Schools
5-point API goal	.4219		
	(.4939)		
>5-point API goal	.3028		
	(.4595)		
<5-point API goal	.2753		
	(.4467)		
Overweight/Obese	31.9836	32.914	31.3046
	(12.2237)	(10.3374)	(13.3939)
Free/Reduced-Price Lunch	54.0933	56.9729	51.9916
	(30.6594)	(23.5266)	(34.8092)
Nonwhite	68.0076	68.5226	67.6316
	(26.8345)	(24.2275)	(28.5825)
English	20.5992	18.7663	21.9371
	(19.1438)	(12.1995)	(22.8282)
Math	23.9437	22.678	24.8676
	(18.5253)	(12.1634)	(21.9928)
API gain last year	10.4467	9.0157	11.4911
	(21.1423)	(20.5933)	(21.4747)
Current API target	-12.9663	3.7254	-25.1494
	(37.7527)	(5.6063)	(45.726)
Observations	27116	11441	15675

Notes: Means (std. deviations) shown. AYP and API variables refer to the previous year, demographic variables to the current year. Observations are at the school-year level.

Table 2:
Effect of Accountability Pressures on the Overweight Status of Students

	(1) Arkansas	(2) Arkansas	(3) California	(4) California
Marginal	0.824**		-0.105	
	(0.276)		(0.154)	
Gain		0.026**		0.010**
		(0.004)		(0.003)
Constant	29.213**	30.033**	28.464**	28.518**
	(3.689)	(3.579)	(1.537)	(1.545)
Observations	2632	2625	27116	27116
R-squared	0.30	0.30	0.39	0.39

Notes: Standard errors (in parentheses) are corrected for within school correlation. All models include a time trend and four powers of the overall literacy rate relative to AYP, four powers of the overall math rate relative to AYP, four powers of the percent of students who are nonwhite, and four powers of the percent of students who are poor. Model (4) also includes two powers of the API target. For model (1) marginal is defined as the lowest scoring subgroup in the school being within 5 points of the AYP goal, while for model (3) marginal is defined as the overall school API goal being 5 points. For model (2) gain is defined as the overall school combined increase in English and math proficiency rates, while for model (4) gain is defined as the overall school increase in API score. California does not use the standard CDC definitions of overweight and obese, but captures the same concept as Arkansas does by using this definition.