You're Not You When You're Hungry: Measuring The Impact of a Weekend Nutrition Program on Childhood Test Scores and Attendance*

Daniel Mangrum†

September 22, 2025

Abstract

This paper examines the impact of a weekend nutrition program on student achievement and attendance in low-income elementary schools in the Mississippi Delta. The intervention provided meals designed to replicate school breakfasts and lunches for weekend consumption. Using a difference-in-differences design, I estimate effects on both the mean and distribution of standardized test scores. Treated students performed better in both language arts and mathematics, with gains of 0.24–0.28 standard deviations in language arts and about 0.16 standard deviations in math. These improvements closed roughly 70–90 percent of the pre-existing test score gap with the state average and were driven by reductions in the share of students at the lowest achievement threshold and corresponding increases at higher levels. Using administrative daily attendance records, I also find significant improvements in attendance: while Friday gains are consistent with the incentive of receiving food bundles, additional increases on Mondays and Tuesdays suggest improved nutrition carried over into the school week. Together, the findings provide causal evidence that weekend feeding programs can meaningfully improve academic outcomes and attendance in food-insecure settings.

Keywords: nutrition, education, test scores, food insecurity.
JEL: I15, I25, I26, I28, H51, H52.

*The opinions expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System. Prior to circulation, this paper was reviewed in accordance with the Federal Reserve Bank of New York review policy, available at `https://www.newyorkfed.org/research/staff_reports/index.html`. The author has no financial interests to disclose.

†Federal Reserve Bank of New York; `daniel.mangrum@ny.frb.org`

1 Introduction

Food insecurity in the United States remains a persistent policy challenge despite decades of federal, state, and local efforts to alleviate hunger for low-income households. In 2023, nearly one in five children experienced at least one episode of food insecurity. Rates vary considerably across regions, with the lowest prevalence in the Midwest and the highest in the South. Within the South, Mississippi has consistently ranked at the top over the last decade, with nearly one in four children classified as food insecure (Dewey et al., 2025). A growing body of research demonstrates that inadequate nutrition undermines children's ability to concentrate and learn, while interventions that reduce food insecurity improve test scores and reduce behavioral problems (Frisvold, 2015; Schwartz and Rothbart, 2020; Figlio and Winicki, 2005; Maluccio et al., 2009; Bond et al., 2022).

This paper examines the impact of a weekend nutrition program that targeted elementary school children from low-income households in the Mississippi Delta, one of the most food-insecure regions in the country. The intervention extended the meals provided through the Free and Reduced Price Lunch Program (FRLP) into the weekend, offering children in grades three through five at two schools food bundles designed to replicate school breakfasts and lunches. Surveys of teachers and parents suggested improvements in attendance and behavior. Here, I provide a more rigorous evaluation using grade-by-school panel data to estimate the program's causal effects on standardized test scores, the distribution of achievement levels, and daily attendance patterns.

Using a difference-in-differences design, I compare treated schools to a set of unaffected schools. I find that the one-time intervention raised student achievement, particularly in language arts, where the average score gap with the state was reduced by nearly 70 percent. Improvements are concentrated among students previously performing at the lowest achievement threshold, with sizable shifts into proficiency. I also find increases in attendance, driven both by higher Friday participation (consistent with the incentive to receive food bundles) and by gains on Mondays and Tuesdays (suggesting improved nutrition and health over the weekend).

This study contributes to the literature on nutrition and education by providing new evidence on weekend feeding programs, a relatively understudied but policy-relevant setting. The results highlight the role of food access in supporting human capital accumulation and point to a cost-effective, easily scalable intervention that could help reduce educational disparities in high-poverty regions.

The remainder of the paper is organized as follows. Section 2 reviews the literature on nutrition interventions to provide benchmark treatment effects. Section 3 discusses the institutional details of this intervention. Section 4 describes the data used in the analysis. Section 5 outlines the empirical strategy. Section 6 presents the main findings, and Section 7 concludes.

2 Literature

The link between nutrition and education has been explored extensively in both developed and developing countries. Many early studies were largely observational or relied on targeting children exhibiting signs of malnutrition. A study of 3,055 third-grade students in Vietnam examined the relationship between anthropometric status and educational achievement and found low test scores in mathematics and Vietnamese were correlated with both low height for age and low weight for age after controlling for age, sex, and school (Hall et al., 2001). A study in Chile surveyed a random sample of children graduating elementary school and high school and the results suggested academic achievement is positively correlated with consumption of dairy, meat, and eggs while consumption of fruits and vegetables is negatively correlated with low academic achievement. Food habits explained nearly 24% of variation in achievement for elementary school students (Ivanovic et al., 1992). A study in the United States used data from the Third National Health and Nutritional Examination Survey to test correlations between food insufficiency and academic achievement and other behavioral outcomes for respondents answering positively to instances of food insufficiency. After controlling for various covariates, the study finds students between 6 and 11 years old were more likely to report not getting along with other students, more likely to repeat a grade, have poorer arithmetic scores, and more likely to have seen a psychologist. Separately studying teenage respondents revealed similar results for behavior responses but no significant results for arithmetic scores suggesting food insecurity is more damaging to younger children than teenagers (Alaimo, Olson, and Frongillo, 2001).

In the case of studies using observation data, omitted variables correlated with diet and academic performance can bias the estimated effect of nutrition on achievement. As a result, more recent studies use more rigorous identification strategies to estimate the casual effect of nutrition on educational achievement. In 1995, the Minnesota state legislature approved a grant providing free breakfast to a set of six treatment schools to test the effectiveness of extending a similar program statewide. In addition to the six treatment schools, three control schools were chosen to provide a natural set of counter-factual students. During the three years of treatment, each treated school reported an increase in math and reading achievement while the control schools' scores were relatively flat. In addition, teachers in treatment schools reported fewer students complained of headaches and stomachaches, students were more energetic, and had an easier time concentrating. Teachers also reported fewer disciplinary issues in the treated schools with morning disciplinary referrals declining between 15 and 50% for all treated schools (Wahlstrom and Begalle, 1999). The use of a set of control schools and the use of administrative data allow for a more plausible argument for causality.

A randomized control trial (RCT) in Jamaica selected seventh grade students in the lower-third of academic performance and randomized the students into a control group and two treatment groups. In one treatment group, students were given a school lunch and in the other treatment group, students were given a syrup drink. The study found that the students who received the school lunch performed better than the control and alternate treatment on an arithmetic test and had better attendance records. The results were also robust to controlling for attendance (Powell, Grantham-McGregor, and Elston, 1983). Another study in Jamaica randomized a breakfast treatment to rural students in grades 2 through 5. Treated students received a school breakfast while control students received one-quarter of an orange and an equal amount of attention. While treated students were shown to have improved in height and weight, significant arithmetic results were only apparent in the youngest of treated children (Powell et al., 1998). An RCT conducted in South Africa randomized 108 students into a treatment and control group and provided breakfast every school day to the treatment group for six weeks. The school breakfast was found to have a positive effect on cognitive performance for the treatment group (Richter, Rose, and Griesel, 1997).

In contrast to randomized control trials, the use of longitudinal data to study diet and academic achievement has become popular due to the ability to perform within-unit comparisons over time. While observational analysis is prone to omitted variable bias, panel data on children can eliminate time invariant unobserved heterogeneity that could otherwise bias observational studies. One particular study performed a randomized control trial and followed Guatemalan children during early childhood and through adulthood. Between 1969 and 1977, four villages in were randomly assigned a high protein drink and a low protein drink meant to be given to children between birth and 36 months of age. The children that were randomly chosen for treatment into the more nutritious drink were tracked and interviewed in 2002. The study found positive effects of the intervention for the treated children. Treated women were found to have completed 1.2 more grades while both treated men and women had increases in both reading comprehension and non-verbal cognitive ability of one-quarter of a standard deviation (Maluccio et al., 2009).

In the case of the Guatemalan study, statistically and economically significant results can be found long after the end of treatment. This suggests that investments made to enhance nutrition in children at critical stages can have lasting effects on educational outcomes. On the other hand, an analysis of standardized test scores in Virginia show that interventions lasting as short as a week can have significant and immediate effects. This study identified schools that were under the threat of accountability sanctions if mean test scores were not improved. Researchers discovered that the school administrators systematically altered school lunch menus in an attempt to increase caloric counts during testing and finds that the schools who increased the caloric content of lunches the most saw the highest test score gains with an increase of 100 calories corresponding to increases of 7, 4, and 7 percentage points for mathematics, English, and social studies, respectively (Figlio and Winicki, 2005).

A more recent wave of studies estimate the effect of the National School Lunch Program and the rollout of School Breakfast programs across the U.S. Frisvold (2015) conducts a rigorous analysis of the School Breakfast Program using two identification strategies and multiple datasets and finds improvements in mathematics of 0.09 standard deviations and improvements in reading of 0.05 standard deviations for schools that adopt the School Breakfast Program. Similarly, Schwartz and Rothbart (2020) studies the impact of universal free lunch in New York City middle schools and the resulting impact on achievement. The study finds that an additional school lunch every two weeks improves math scores by around 0.08 standard deviations and improves language arts test scores by around 0.07 standard deviations. On the contrary, one recent study finds evidence of lower test scores when students are furthest away from the benefit receipt date. The study finds that when the students' family received SNAP benefits between 27 and 30 days prior to the test, math scores decline on average between 0.024 and 0.046 standard deviations (Cotti, Gordanier, and Ozturk, 2018). This suggests that families who receive federal food benefits may exhibit food insecurity at the end of the benefits cycle and that food insecurity may adversely impact academic performance.

While the literature covers a multitude of interventions spanning nutrition supplements, food stamps, school lunches, and school breakfasts, there is little evidence in the literature covering interventions providing supplemental nutrition over the weekend for students reliant on free and reduced school lunches during the week. Since the introduction of the School Lunch Program and Free or Reduced Lunch Program, families have become increasingly reliant on school lunches to provide adequate nutrition to children in low socioeconomic settings. In addition to school lunches, schools with a large proportion of free or reduced lunch eligible students tend to also provide breakfast to students. While these students receive a majority of their caloric intake from school during the school-week, the students must rely on their own household for nutrition over the weekend.

The most similar study to this paper examining a weekend nutrition intervention for elementary students in the U.S. is Kurtz, Conway, and Mohr (2020), which analyzes the roll-out of a "backpacks" nutrition program across counties in North Carolina between 2009 and 2013. Their study documents meaningful improvements in attendance, particularly on Fridays, and gains in test scores, underscoring the potential of weekend feeding programs to improve student outcomes. At the same time, their evaluation focuses on a *targeted* intervention in which only certain students at participating schools were selected for treatment, while outcomes were measured for both selected and non-selected students. By contrast, the setting I study involves an un-targeted intervention in which all students in grades 3 through 5 at the treated schools were offered inclusion, and more than 80% returned permission forms to participate. This universal approach likely mitigates potential stigma associated with participation (Fram and Frongillo, 2018) and means that the Intent to Treat (ITT) estimates in this paper are closer to the Average Treatment Effect on the Treated (ATT), since the vast majority of eligible students were included. In this way, this study complements and extends prior evidence by providing causal estimates from a setting where the intervention reached nearly the entire eligible population.

Table 1: Contents of Weekend Supplemental Nutrition

2 bowls of cereal
2 small containers of fat-free milk
2 pieces fresh fruit
3 canned meats
1 cup applesauce
1 cup mixed fruit in syrup

3 Background

During the 2011-2012 school year, a local non-profit conducted an intervention in two elementary schools in the Mississippi Delta. Third- and fourth-grade students at Brooks Elementary School in Bolivar County and third-, fourth-, and fifth-grade students at Stampley Elementary School in Coahoma County were selected for inclusion into the treatment. Students at these schools were overwhelmingly eligible for the Free and Reduced Lunch Program (FRLP) with 99% eligibility at Brooks Elementary and 94% eligibility at Stampley Elementary. According to questionnaires administered by the non-profit, food insecurity was prevalent at both schools with more than half of parents at the two schools reporting some degree of food insecurity.

In order to be included in the intervention, students were required to return a permission form signed by a parent or guardian. In total, 174 students were included in the treatment (73 at Stampley and 101 at Brooks).[1] Each Friday between September 30, 2011 and May 18, 2012 recyclable bags filled with food were distributed to all students involved in the treatment. Each bag contained food intended to last for weekend consumption denoted in Table 1. In order for the student to receive the food, the student must have attended school on Friday. Food was not distributed over weeks in which students were on break including: Fall Break, Thanksgiving Break, Winter Break, and Spring Break.

4 Data

I use data from the Mississippi Department of Education that include performance metrics on the Mississippi Curriculum Test, 2nd Edition (MCT2). The MCT2 is the standardized test

[1] In 2011-2012, there were 56 third-graders and 51 fourth-graders who took the standardized tests at Brooks Elementary. At Stampley, there were 28 third-graders, 34 fourth-graders, and 38 fifth-graders who took the tests.

administered every spring to public school students in the state of Mississippi during the sample period. MCT2 tests students in Language Arts and Mathematics beginning in the third grade. The data for this analysis spans the academic years 2007-2008 through the 2011-2012 school years. MCT2 scores are reported for each public school in the state of Mississippi for each grade level. In addition to average scores, the data also include information about the distribution of test scores. This data include the percentage of students in each school-grade cell achieving between three thresholds creating four bins of test scores: Minimal, Basic, Proficient, and Advanced. These metrics allow for analysis on both the average and the distributional change in test scores as a result of the intervention.

I also obtained attendance records for all Mississippi public schools for the sample years. These records contain the number of students absent from each school-grade cell on a daily basis throughout each school year. To create an outcome measure that is consistent across various school sizes, I compute the number of absences by day-of-week per enrolled student at the school-grade cell. Enrollment data comes from the Common Code of Data (CCD).

Lastly, I include control variables for changes in other programs that might also affect food insecurity. I merge data on Free and Reduced Lunch Program (FRLP) eligibility for each school-grade-year cell from the Common Core of Data and I include county level Supplemental Nutrition Assistance Program (enrollment) for each school. Table 2 reports summary statistics for outcome measures and demographics between the 2007-2008 and 2010-2011 school years.[2] Column 1 includes the full sample of Mississippi elementary schools with a highest grade of five. Column 2 contains only the two treated schools, J.W. Stampley Elementary and Brooks Elementary. Comparing these columns reveals that the schools selected for the intervention score worse than the average Mississippi elementary school in both Language Arts and Mathematics test scores. The distributions for both tests follow a similar pattern with the treated schools having a larger percentage of students achieving Minimal and Basic standards and fewer students achieving Proficient and Advanced. Students at the schools selected for treatment also missed more days of school per student with those increases largely occurring on Mondays and Fridays. Lastly, the treated schools are populated by a larger share of Black students, have smaller average grade sizes, and

[2] Schools with missing information on the percent of students receiving Free and Reduced School lunches are omitted from the analysis.

Table 2: Descriptive Statistics: Test Scores and Demographic Variables

	(1)	(2)
Variable	All MS Schools	Treated Schools
LA Score	–0.056	–0.317
% Minimal	15.6	20.1
% Basic	36.4	45.5
% Proficient	37.4	27.9
% Advanced	10.6	6.4
Test Takers	91.4	42.7
Math Score	–0.053	–0.245
% Minimal	16.5	16.3
% Basic	28.5	37.9
% Proficient	44.9	41.6
% Advanced	10.1	4.3
Test Takers	91.4	42.7
Absences per student	5.97	6.88
Monday	1.31	1.53
Tuesday	1.16	1.32
Wednesday	1.10	1.33
Thursday	1.08	1.19
Friday	1.32	1.50
% White	38.3	0.0
% Black	57.4	98.4
% Hispanic	2.1	0.3
% Asian	0.5	0.0
% Free Lunch	79.2	98.3
Schools	498	2

The table above reports means for outcome variables (test scores) and demographic variables for the full sample of MS schools separately from the two treated schools. Outcome variables include normalized mean test scores in Language Arts and Mathematics on the MCT2 state test along with the number of test takers for each school-grade level. Demographic variables include the proportion of each school-grade that is white, black, Hispanic, or Asian as well as the percent of each school-grade cell that is eligible for Free and Reduced School Lunch. Demographic and FRLP data come from IPEDS while test score data come from Mississippi Department of Education.

have higher FRLP eligibility.

5 Empirical Strategy

5.1 Identification

Previous qualitative surveys administered to teachers and parents of the treated students suggest that students were better behaved, more attentive in class, and more likely to attend school on Friday during the year of the intervention.[3] However, these impacts cannot be interpreted as casual effects of the intervention without considering how other students performed during the 2011-2012 school year. It is possible students that were not exposed to the intervention also performed better in school and had better attendance in the 2011-2012 school year for reasons unrelated to the intervention. Further, it could be the case that parents and teachers were more likely to answer positively on these surveys as a result of knowledge of the treatment in hopes of continued aid.

To help mitigate the concerns, I employ a difference-in-differences identification strategy to estimate the effect of the supplemental nutrition program on test scores and attendance for the students in the schools selected for the intervention. This design compares changes in outcomes of treated school-grade units to changes in outcomes for otherwise similar untreated school-grade units to control for any factors that might influence students in the same manner in the same years. If the necessary assumption, discussed below, is satisfied the estimated effect is the Intent to Treat (ITT) estimate of the intervention.[4]

The baseline specification for the test score outcomes follows a relatively standard difference-in-differences framework:

$$y_{gst} = \gamma treat_{st} + \delta_g + \delta_s + \delta_t + \delta_s \times t + \beta X_{gst} + \varepsilon_{gst}, \tag{1}$$

where y_{gst} is an outcome measure for grade g at school s during academic year t. $treat_{st}$ is a binary variable which equals one for the two treated schools in the 2011-2012 academic year where γ is

[3]From summary report from the administering non-profit organization. Available upon request.

[4]Since not all students returned permission slips to enroll in the intervention, the effect estimated is not the Average Treatment Effect on the Treated.

the estimated ITT. Also included are fixed effects for schools (δ_s), grade levels (δ_g), and academic years (δ_t). I also include a school specific linear time trend to capture differences in test score trends for each school. As mentioned above, I include time varying controls for the share of students eligible for FRPL, county unemployment, and county enrollment in SNAP.

Typical cluster robust standard errors perform remarkably poorly in difference-in-difference settings with few treated clusters (MacKinnon and Webb, 2018; Cameron, Gelbach, and Miller, 2008; Cameron and Miller, 2015). Hence, in our setting with only two treated clusters, where the clusters are schools, typically computed standard errors will over-reject the null hypothesis. Instead, I report p-values from Wild Cluster Bootstrap with 9,999 replications, clustered at the school level, which are more conservative and are less likely to over-reject the null hypothesis (Djogbenou, MacKinnon, and Nielsen, 2019; MacKinnon and Webb, 2018).

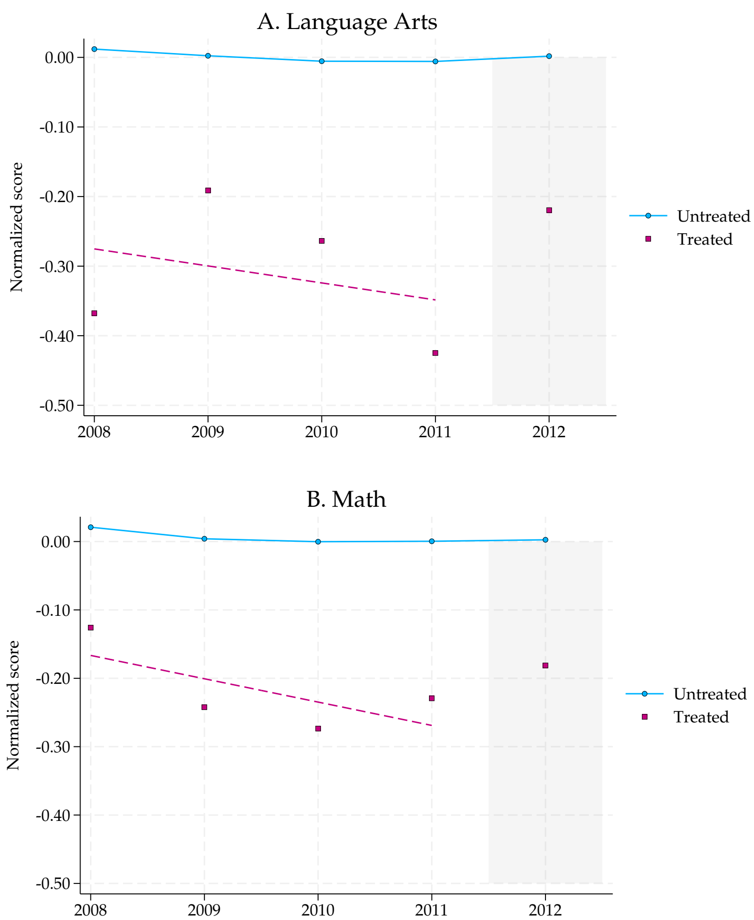
5.2 Selection of Comparison Group

As mentioned above, the estimated coefficient γ can only be interpreted as the ITT if the treated units would have evolved similarly to the control units in the absence of treatment. If the untreated units used as controls in the sample are systematically different from the treated units, the observed outcomes in the year of treatment cannot be reasonably used as a counter-factual outcome for the treated units. Thus, it must be the case that *changes* in test scores in the academic years prior to 2011-2012 are similar across the treated and untreated units. As described in Section 4, the schools selected for treatment had, on average, lower test scores, higher participation in FRLP, and had a larger share of Black students than the overall sample of schools in Mississippi. Figure 1 shows the average trend in Language Arts and Mathematics test scores for all grades in the schools selected for treatment versus the sample of all other Mississippi Schools. Since the test scores for the full sample are normalized, the full sample of schools will be (near) mean zero for all years.[5] However, Language Arts and Math scores for the treated schools are around 0.3 and 0.2 standard deviations lower than the mean, respectively. Additionally, average scores are slightly decreasing in the years leading up to the intervention. This trend could be the result of falling raw scores in the schools selected for treatment, gains in other schools, or both. Regardless, the trend in scores between the schools selected for treatment and the full sample of schools suggests that the full sample of

[5]Means are not exactly zero because some schools are removed from the sample for missing data.

schools might make for a suitable control group for the treated schools despite the level differences if the linear trend is accounted for.

Figure 1: Mean Test Score Trends for Full Sample versus Treated Schools



Each figure above plots the mean test scores for the full sample of MS schools against the mean test scores for the treated schools in each year of the data. In addition, the red dashed line shows the trend in test scores for the treated schools in the three years before treatment. The shaded area, the 2011-2012 school year, denotes the year of treatment.

I also select a subsample of schools from the sample of all Mississippi public schools that is more similar to the two treated schools on the observables in which the treated schools most differ from the full sample: FRLP eligibility, size of grade level, and proportion of the student body who is Black. In addition, I also consider distance from the two treated schools as-the-crow-flies. In many cases, matching (or inverse probability weighting) methods use a nonlinear model such as a probit or a logit model to find comparison units that are similar to the units selected for treatment. However, in this setting, a nonlinear model is not feasible since there are few schools selected for treatment and because the treated schools have extreme values for the matching variables. Instead, I use Mahalanobis multivariate distance (MD) to rank the schools on similarity metrics. This method allocates more weight to schools that are more similar to the two treatment schools (Rubin, 1980). The algorithm takes selection variables and corresponding values on which to the match observations and calculates the distance between each observation in the sample from the mean values of the treated schools. The distance for each matching variable is weighted by the inverse of the variance-covariance matrix of the selection variables. The result is a weighted distance measure, measured in standard deviations, ranking how similar units are to the variable values on which the matching is constructed.

Using this similarity measure, I select a subsample of most similar schools by choosing the top quartile of schools based on Mahalanobis distance. Table 3 details the 13 most similar schools alongside the two treated schools as well as the 15 least similar schools among the top quartile. Among the most similar schools, all are 95% or more FRLP participation, 88% or more in the Black population of the school, and have between 29 and 65 test takers. None are more than 34 miles on average from the treated schools. The most dissimilar schools within the matched subsample tend to match the treated schools on most dimensions but differ along one or two. For instance, Woodville Heights Elementary has 89% FRLP participation and is 97% Black students, however class size is much larger at 88 test takers per grade and the school is 129 miles from the treated school on average. Despite the larger class size and the distance, this is likely still a comparable school to be considered as a control unit, as are the remaining least similar schools in the upper quartile of similarity.

Next, we test to see whether matching on demographics using Mahalanobis distance creates a comparison group that better matches outcomes of the treated schools on levels and trends. Table 4 repeats Table 2 adding the summary statistics for the schools in the comparison group determined by the constructed Mahalanobis multivariate distance measure. Selection on the three demographic characteristics and the average distance creates a comparison group that is much more similar to the treated units than the overall sample. The 165 schools in the comparison group only differ in FRLP eligibility by 5.8 percentage points and differ in the proportion of Black students by only 13.1 percentage points compared to the overall sample difference of 19.1 and 41 percentage points, respectively. The number of test takers for each grade differs by 22 students compared

Table 3: Comparison Group by Mahalanobis Multivariate Distance

Rank	School	District	% FRLP	% Black	Test Takers	Avg. Dist	Mahalanobis Distance
1	Shelby School	North Bolivar	98.6	98.2	47.5	12.8	0.052
2	I T Montgomery Elementary	Mound Bayou	99.0	99.2	44.3	17.2	0.076
3	Lyon Elementary	Coahoma County	98.9	97.6	40.4	11.7	0.165
4	Myrtle Hall IV Elementary	Clarksdale	98.4	99.6	47.9	8.2	0.174
5	**J W Stamply Elementary**	Clarksdale	98.1	98.6	36.4	0.0	0.192
6	**Brooks Elementary**	North Bolivar	98.9	97.4	52.6	0.0	0.192
7	A W James Elementary	Drew	96.1	91.8	49.2	23.5	0.216
8	Nailor Elementary	Cleveland	98.7	97.0	33.3	26.6	0.237
9	Booker T Washington	Clarksdale	96.8	99.0	43.9	7.7	0.244
10	Hunter Middle	Drew	96.4	96.3	41.3	23.7	0.257
11	Jonestown Elementary	Coahoma County	98.9	99.6	44.4	18.3	0.300
12	McEvans School	Shaw	97.0	96.8	45.0	33.8	0.321
13	West Bolivar Elementary	West Bolivar	96.2	92.4	64.4	28.5	0.330
14	Friars Point Elementary	Coahoma County	98.0	97.6	29.5	17.5	0.332
15	Kirkpatrick School	Clarksdale	96.1	88.2	46.7	7.6	0.336
153	Eupora Elementary	Webster County	73.0	31.6	75.7	89.0	2.014
154	Louisville Elementary	Louisville	91.6	82.2	143.7	115.6	2.022
155	Senatobia Middle	Senatobia	61.8	45.0	143.6	52.5	2.039
156	Hawkins Middle	Forest City	86.3	68.2	124.2	139.7	2.044
157	North Jackson Elementary	Jackson Public	88.3	98.8	76.3	123.0	2.070
158	Beechwood Elementary	Vicksburg-Warren	67.8	47.4	68.4	123.6	2.077
159	Woodville Heights Elementary	Jackson Public	89.3	97.2	87.9	129.4	2.079
160	Stokes Beard Elementary	Columbus	91.3	93.3	61.0	137.0	2.080
161	Houston Upper Elementary	Houston Separate	76.1	43.0	148.3	96.2	2.090
162	Hope Sullivan Elementary	DeSoto County	64.8	31.5	136.8	69.8	2.122
163	Lake Middle	Scott County	75.0	39.4	55.6	139.8	2.126
164	Vicksburg Interm	Vicksburg-Warren	90.7	79.4	150.1	127.8	2.129
165	Parks Elementary	Cleveland	70.8	30.6	55.8	26.9	2.137
166	Greenlee Attendance Center	Attala County	78.0	31.0	51.7	97.2	2.137
167	Forest Elementary	Forest City	89.6	59.2	118.2	138.5	2.138

The table above ranks the top 15 and bottom 15 schools selected in the subsample as determined by the Mahalanobis multivariate distance using the mean values of the two treated schools for %FRLP, % Black, number of Test Takers and average distance from Brooks and Stampley Elementary. Data on FRLP eligibility and % Black come from IPEDS while the number of test takers comes from the Mississippi Department of Education. Avg. Dist. is the average distance between the school and Brooks Elementary and Stampley Elementary and are computed using the Stata program geodist which calculates distance as-the-crow-flies. Treated schools in bold.

to the overall sample difference of 48.7. The more suitable control group generates average test scores that are also far more similar than the overall sample: average test scores in Language Arts are within 0.057 standard deviations and test scores in Math are within 0.007 standard deviations.

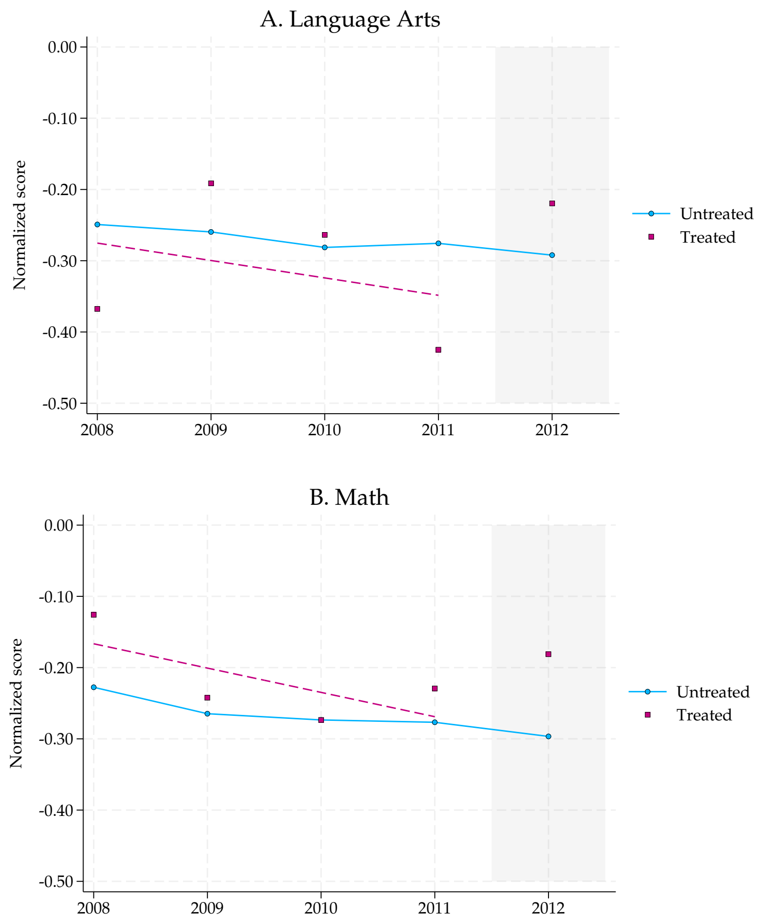
Figure 2 repeats Figure 1 with the comparison group selected using the Mahalanobis similarity score. This subsample more closely matches the *level* of test scores for both Language Arts and Math than the full sample and a better job matching the *trend* in scores. In this figure, the potential treatment impact is more apparent with 2012 Language Arts and Math scores for the treated units outpacing the same year scores for the comparison group. This trend break suggests that the intervention had a positive impact on the treated schools in both subject areas. If the 2012 scores for the comparison group are used as a counter-factual for the treated units, this would suggest improvements in test scores larger than 0.1 standard deviations. In the next section, I present results from the estimation of Equation (1) for the full sample and the selected comparison group in order to estimate the causal impact of the supplemental nutrition intervention on test scores and attendance.

Table 4: Descriptive Statistics: Test Scores and Demographic Variables

	(1)	(2)	(3)
Variable	All MS Schools	MD Matched Schools	Treated Schools
LA Score	–0.056	–0.260	–0.317
% Minimal	15.6	20.0	20.1
% Basic	36.4	41.4	45.5
% Proficient	37.4	31.9	27.9
% Advanced	10.6	6.7	6.4
Test Takers	91.4	64.7	42.7
Math Score	–0.053	–0.252	–0.245
% Minimal	16.5	21.1	16.3
% Basic	28.5	32.4	37.9
% Proficient	44.9	40.6	41.6
% Advanced	10.1	5.9	4.3
Test Takers	91.4	64.7	42.7
Absences per student	5.97	5.67	6.88
Monday	1.31	1.25	1.53
Tuesday	1.16	1.09	1.32
Wednesday	1.10	1.04	1.33
Thursday	1.08	1.02	1.19
Friday	1.32	1.26	1.50
% White	38.3	11.8	0.0
% Black	57.4	85.3	98.4
% Hispanic	2.1	1.4	0.3
% Asian	0.5	0.1	0.0
% Free Lunch	79.2	92.5	98.3
Schools	498	165	2

The table above reports means for outcome variables (test scores) and demographic variables for the full sample of MS schools, the upper quartile of MS schools most similar to the treated schools by Mahalanobis multivariate distance, and the two treated schools. Outcome variables include normalized mean test scores in Language Arts and Mathematics on the MCT2 state test along with the number of test takers for each school-grade level. Demographic variables include the proportion of each school-grade that is white, black, Hispanic, or Asian as well as the percent of each school-grade cell that is eligible for Free and Reduced School Lunch. Demographic and FRLP data come from IPEDS while test score data come from Mississippi Department of Education.

Figure 2: Mean Test Score Trends for Matched Schools versus Treated Schools



Each figure above plots the mean test scores for the subsample of MS schools most similar to the treated group as described in 5 against the mean test scores for the treated schools in each year of the data. In addition, the red dashed line shows the trend in test scores for the treated schools in the three years before treatment. The shaded area, the 2011-2012 school year, denotes the year of treatment.

6 Results

6.1 Standardized Test Scores

Table 5 reports the results from the estimation of Equation (1). Columns 1 and 2 report the results for Language Arts scores while columns 3 through 4 report the results for Math scores. The first column for each subject score includes the full sample while the second column includes only the subsample of the most similar quartile of schools based on the multivariate distance measure described above. For each outcome, I report Wild Cluster Bootstrap p-values clustered at the school level to empirically test whether the estimate is statistically different from zero. The first two columns suggest improvements in Language Arts test scores regardless of the comparison group. The estimates range from 0.24 to 0.28 from a pre-treatment mean of -0.317. Hence, the results suggest the intervention closed 77% to 88% of the gap in test scores between the treated schools and the state average. The effect on math scores is slightly smaller at around 0.16 standard deviation improvement from a mean of -0.245 representing a closing of the test score gap of 67%. In total, the point estimates suggest significant test score improvements as a results of the nutrition intervention for the school-grade cells selected for treatment. Using the Wild Cluster Bootstrap, which allows for efficiency gains over typical cluster robust standard errors when there are few treated clusters, results in p-values less than 0.01 in each specification and for each outcome.

Table 5: Language Arts and Math Scores

	Language Arts Score		Math Score	
	(1)	(2)	(3)	(4)
Treatment	0.244***	0.278***	0.165***	0.167***
	[0.000]	[0.000]	[0.000]	[0.000]
Outcome Mean	–0.317	–0.317	–0.245	–0.245
Observations	5,787	2,079	5,787	2,079
Schools	499	167	499	167
Sample	Full	Matched	Full	Matched

Note: Each column contains the estimated difference-in-differences coefficient where the outcome variable is either the standardized language arts or mathematics test score as denoted in the column title for a school-grade cell. Columns One and Three use the full sample of schools. Columns Two and Four use the matched subsample of schools using the Mahalanobis distance measure. The Outcome Mean reports the mean value for the given outcome across the treated schools in the years prior to treatment. P-values using 9,999 wild cluster bootstrap (WCB) replications are presented in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6.2 Distributional Treatment Effects

The discussion in the previous section presents evidence that average standardized test scores for the schools selected for treatment improved in the year of the intervention. The magnitude of these estimates is quite large and suggest that the treated students performed almost as well as the average Mississippi student despite having average test scores nearly a third of a standard deviation below average in previous years. To further explore how these test scores shifted, we explore movements in the shares of students meeting certain achievement thresholds. Table 6 presents the difference-in-difference estimates for the percentage of students in each performance bin from worst (Minimal) in Panel A to best (Advanced) in Panel D for both language arts and

Table 6: Distributional Effects for Language Arts and Mathematics Test Scores

	Language Arts Score		Math Score	
A. Percent Minimal	(1)	(2)	(3)	(4)
Treatment	–8.5*	–9.0***	–6.5	–6.6
	[0.064]	[0.001]	[0.391]	[0.265]
Outcome Mean	20.147	20.147	16.284	16.284
B. Percent Basic				
Treatment	–11.6	–10.8	3.7***	4.2***
	[0.287]	[0.163]	[0.000]	[0.000]
Outcome Mean	45.511	45.511	37.858	37.858
C. Percent Proficient				
Treatment	19.1	17.9	–2.6	–3.4
	[0.384]	[0.391]	[0.504]	[0.465]
Outcome Mean	27.884	27.884	41.595	41.595
D. Percent Advanced				
Treatment	1.0	2.0	5.5***	5.8***
	[0.499]	[0.289]	[0.000]	[0.000]
Outcome Mean	6.447	6.447	4.268	4.268
Observations	5,787	2,079	5,787	2,079
Schools	499	167	499	167
Sample	Full	Matched	Full	Matched

Note: Each column contains the estimated difference-in-differences coefficient. Each panel contains a different outcome variable where the outcome is the percentage of students in the school-grade cell achieving at a certain achievement threshold. The thresholds increase in achievement level from Minimal (lowest), Basic, Proficient, and Advanced (highest). Columns One and Three use the full sample of schools. Columns Two and Four use the matched subsample of schools using the Mahalanobis distance measure. The Outcome Mean reports the mean value for the given outcome across the treated schools in the years prior to treatment. P-values using 9,999 wild cluster bootstrap (WCB) replications are presented in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

mathematics. For both language arts and mathematics, there are declines in the percentage of students achieving at the lower threshold. The effect is larger again for language arts and the point estimate suggests a 42% improvement.[6] The reduction for mathematics is similar in proportional terms but smaller in magnitude. However, the reduction in students achieving the Minimal standard is statistically difference from zero only for Language Arts.

The results for the second bin of achievement are mixed across language arts and mathematics. For language arts, the point estimates suggest declines around 11 percentage points while math achievement in this bin is around 4 percentage points larger. In contrast to the Minimal standard, the reduction in Basic achievement for language arts is not statistically different from zero (p-value

[6] 8.5 percentage point reduction off of a base of 20.1 percentage points

equal to 0.16 for the matched sample), the increase in Basic achievement for math is precisely estimated for the treated schools.

Since students move from one bin to another by construction, the decline in both Minimal and Basic for language arts suggest students are shifting into higher achievement. This result is reflected in the point estimates in Panel C for the percentage of students achieving Proficient standards. The point estimate is quite large at between 17.9 and 19.1 percentage points, but neither estimate is statistically significant. However, this still represents suggestive evidence that treated students are shifting from lower achievement bins to higher achievement bins for language arts scores. On the contrary, the point estimates suggest a reduction in students achieving Proficient in math. This could be due to students leaving the Proficient bin and moving into either Basic or Advanced. The results in Panel D for mathematics hint that the latter is more plausible. While the estimates for language art are economically small and are imprecisely estimated, the effect for mathematics is quite large and precisely estimated relative to the baseline average for the treated schools in previous years. This large improvement in mathematics might explain the negative point estimates for the Proficient bin in Panel C: it may be the case that students previously achieving at Minimal are able to move to Basic and some students who previously achieved at Proficient are able to move into Advanced.

In total, analysis on the distribution of test scores suggests reductions in students achieving at lower thresholds with shifts towards higher achievement bins. For language arts, these improvements largely stem from students exiting the Minimal and Basic bins and entering the Proficient bins. For mathematics, there is some evidence students leave the Minimal bin and other students enter the Advanced bin but without student-level data, it is not possible to track which students move across which bins.

6.3 Attendance

In addition to test score outcomes, I also explore how daily attendance is affected by the supplemental nutrition intervention. The results from this section can help to explain the mechanisms by which the nutrition intervention improves test scores. If the supplemental nutrition program indeed improves nutrition over the weekend, students should feel healthier and be less likely to

Table 7: Absences per Enrolled Student by Day-of-Week

A. Overall	(1)	(2)
Treatment	–1.205	–1.251
	[0.342]	[0.149]
Outcome Mean	6.240	6.240
B. Monday		
Treatment	–0.318	–0.320
	[0.494]	[0.470]
Outcome Mean	1.418	1.418
C. Tuesday		
Treatment	–0.248***	–0.288***
	[0.000]	[0.000]
Outcome Mean	1.214	1.214
D. Wednesday		
Treatment	0.040	0.019
	[0.679]	[0.848]
Outcome Mean	1.168	1.168
E. Thursday		
Treatment	–0.046	–0.052
	[0.676]	[0.657]
Outcome Mean	1.074	1.074
F. Friday		
Treatment	–0.632**	–0.610***
	[0.011]	[0.001]
Outcome Mean	1.366	1.366
Observations	5,790	2,081
Schools	500	168
Sample	Full	Matched

Note: Each column contains the estimated difference-in-differences coefficient where the outcome variable is the number of absences on each day of the week per enrolled student. Column One uses the full sample of schools and column Two uses the matched subsample of schools using the Mahalanobis distance measure. The Outcome Mean reports the mean value for the given outcome across the treated schools in the years prior to treatment. P-values using 9,999 wild cluster bootstrap (WCB) replications are presented in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

be absent earlier in the week. Additionally, the intervention acts as an in-kind transfer in which students must attend school on Fridays to receive the transfer. As such, attendance on Fridays should improve.

Table 7 reports the results for the total number of days missed each school year per enrolled student at the school-grade level. I first report for all days before breaking out the absence rate by day of week. Panel A reports the results for the total absence rate. Both columns report similar estimates, a reduction of around 1.2 days missed per enrolled student, however the estimates are not statistically difference from zero – the p-value on the matched sample is 0.149. Panel B reports these estimates for Mondays. On average, students in the treated schools missed around 1.4 Mondays per year making Monday the most frequently missed school day of the week. The estimates suggest a reduction in Mondays missed in the year of treatment by 0.32 days per enrolled student representing a 23% decline in Monday absenteeism, however this estimate is very imprecise. The results in Panel C for Tuesdays are similar to those in Panel B with declines in attendance around .29 days per enrolled students, a reduction of 24% from the treated schools' pre-treatment average. The estimate for Tuesdays is also highly significant with a p-value less than 0.01.

Panels D and E suggest no change in attendance for Wednesday or Thursdays. Both estimates are economically small and not precisely estimated across the comparison groups.Lastly, the effect on Friday attendance seems to be the largest across the week, resulting in a reduction of 0.61 missed Fridays during the year per enrolled student. This reduction corresponds to a 45% reduction in Friday absenteeism as a result of the treatment. These results lend evidence to the mechanisms presented above. The largest effect of the intervention on attendance is due to the transfer effect on Fridays, causing almost 50% reductions in absenteeism. However, the evidence also suggests improvements in attendance on Mondays and Tuesday which would likely reflect improvements in health in nutrition over the course of the weekend.

7 Conclusion

This study provides causal evidence that extending school-based nutrition into weekends can substantially improve student outcomes in one of the most food-insecure regions of the United

States. The intervention closed roughly 70 percent of the test score gap in language arts relative to the state average and generated large gains in mathematics as well. Attendance improved markedly, with nearly a 45 percent reduction in Friday absenteeism and measurable gains at the start of the school week. These patterns point to dual mechanisms: a direct incentive for students to attend on Fridays in order to receive food bundles, and improved nutrition over the weekend that enhanced students' readiness to learn on Mondays and Tuesdays.

The benefits of the program were especially concentrated among students at the bottom of the achievement distribution. In language arts, large reductions in the share of students performing at the lowest levels were accompanied by sizable gains in proficiency. For mathematics, some students advanced out of the lowest bin, while others shifted from proficiency to advanced performance, suggesting that the program supported both foundational catch-up and higher-order mastery.

These effects were achieved at very low cost. Using the Kraft (2020) benchmark, the intervention clearly qualifies as a "large effect size/low-cost" program, with weekly food bundles costing a fraction of the $500 per pupil threshold. Because schools already maintain infrastructure to order and distribute meals, extending this capacity to cover weekends requires relatively little new investment. As such, this program represents a scalable and administratively feasible way to deliver large returns in both academic achievement and attendance per dollar spent.

The external validity of these findings is necessarily tied to context: effects are likely largest where food insecurity is most acute. Still, the results highlight nutrition policy as an underutilized lever in education policy. Targeted weekend feeding programs can not only reduce hunger but also help close achievement gaps, especially in persistently disadvantaged regions such as the Mississippi Delta. Scaling such programs within similarly high-need areas offers a cost-effective and equitable strategy for strengthening human capital and promoting educational opportunity.

References

Alaimo, Katherine, Christine M Olson, and Edward A Frongillo. 2001. "Food insufficiency and American school-aged children's cognitive, academic, and psychosocial development." *Pediatrics* 108 (1):44–53.

Bond, Timothy N, Jillian B Carr, Analisa Packham, and Jonathan Smith. 2022. "Hungry for success? SNAP timing, high-stakes exam performance, and college attendance." *American Economic Journal: Economic Policy* 14 (4):51–79.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3):414–427.

Cameron, A. Colin and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2):317–372.

Cotti, Chad, John Gordanier, and Orgul Ozturk. 2018. "When does it count? The timing of food stamp receipt and educational performance." *Economics of Education Review* 66:40–50.

Dewey, Adam, Julie Hilvers, Sena Dawes, Virginia Harris, Monica Hake, and Emily Engelhard. 2025. "Map the Meal Gap: A Report on Local Food Insecurity and Food Costs in the United States in 2023." Technical report, Feeding America National Organization. URL `https://www.feedingamerica.org/sites/default/files/2025-05/Map%20the%20Meal%20Gap%202025%20Report.pdf`. Accessed: 2025-09-16.

Djogbenou, A. A., James G. MacKinnon, and Morten Orregaard Nielsen. 2019. "Asymptotic Theory and Wild Bootstrap Inference with Clustered Errors." *Journal of Econometrics* 212 (2):393–412.

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

Fram, Maryah S and Edward A Frongillo. 2018. "Backpack programs and the crisis narrative of child hunger—a critical review of the rationale, targeting, and potential benefits and harms of an expanding but untested model of practice." *Advances in Nutrition* 9 (1):1–8.

Frisvold, David E. 2015. "Nutrition and cognitive achievement: An evaluation of the School Breakfast Program." *Journal of public economics* 124:91–104.

Hall, Andrew, LN Khanh, TH Son, NQ Dung, RG Lansdown, DT Dar, NT Hanh, Helen Moestue, Ha Huy Khoi, and DA Bundy. 2001. "An association between chronic undernutrition and educational test scores in Vietnamese children." *European Journal of Clinical Nutrition* 55 (9):801–804.

Ivanovic, Daniza, Magaly Vásquez, Marcela Aguayo, Digna Ballester, Maximiliano Marambio, and Isabel Zacarías. 1992. "Nutrition and education. III. Educational achievement and food habits of Chilean elementary and high school graduates." *Archivos latinoamericanos de nutricion* 42 (1):9–14.

Kraft, Matthew A. 2020. "Interpreting effect sizes of education interventions." *Educational researcher* 49 (4):241–253.

Kurtz, Michael D, Karen Smith Conway, and Robert D Mohr. 2020. "Weekend feeding ("BackPack") programs and student outcomes." *Economics of Education Review* 79:102040.

MacKinnon, James G. and Matthew D. Webb. 2018. "The Wild Bootstrap for Few (Treated) Clusters." *The Econometrics Journal* 21 (2):114–135.

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein. 2009. "The impact of improving nutrition during early childhood on education among Guatemalan adults." *The Economic Journal* 119 (537):734–763.

Powell, Christine, Sally Grantham-McGregor, and M Elston. 1983. "An evaluation of giving the Jamaican government school meal to a class of children." *Human Nutrition. Clinical Nutrition* 37 (5):381–388.

Powell, Christine A, Susan P Walker, Susan M Chang, and Sally M Grantham-McGregor. 1998. "Nutrition and education: a randomized trial of the effects of breakfast in rural primary school children." *The American journal of clinical nutrition* 68 (4):873–879.

Richter, Linda M, Cynthia Rose, and R Dev Griesel. 1997. "Cognitive and behavioural effects of a school breakfast." *South African Medical Journal* 87 (1 Suppl):93–100.

Rubin, Donald B. 1980. “Bias reduction using Mahalanobis-metric matching.” *Biometrics* :293–298.

Schwartz, Amy Ellen and Michah W Rothbart. 2020. “Let them eat lunch: The impact of universal free meals on student performance.” *Journal of Policy Analysis and Management* 39 (2):376–410.

Wahlstrom, Kyla L and Mary S Begalle. 1999. “More Than Test Scores: Results of the Universal School Breakfast Pilot in Minnesota.” *Topics in Clinical Nutrition* 15 (1):17–29.