The Effect of Healthy School Lunch Provision on Academic Test Scores

Michael L. Anderson, University of California, Berkeley, mlanderson@berkeley.edu

Justin Gallagher, Case Western Reserve University, justin.p.gallagher@case.edu

Elizabeth Ramirez, University of California, Berkeley, elizabeth.ramirez@berkeley.edu

Selected Paper prepared for presentation at the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meetings, San Francisco, CA, July 26-28

The Effect of Healthy School Lunch Provision on Academic Test Scores

Michael L. Anderson, Justin Gallagher, and Elizabeth Ramirez*

May 22, 2015

*Anderson: 207 Giannini Hall, MC 3310, University of California, Berkeley, CA 94720-3310 (email: mlanderson@berkeley.edu); Gallagher: Department of Economics, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7235 (email: justin.p.gallagher@case.edu); Ramirez: 207 Giannini Hall, MC 3310, University of California, Berkeley, CA 94720-3310 (email: elizabeth.ramirez@berkeley.edu). The authors would like to thank Anthony Gatti, Sarah Mattson, Jonathon Mobley, and Aaron Weisberg for outstanding research assistance. This work was supported by the Giannini Foundation of Agricultural Economics and the USDA National Institute of Food and Agriculture, Hatch Project 233535.

1 Introduction

Improving the nutritional content of school meals is a topic of intense policy interest. The primary motivation underlying these nutritional improvements is to increase student health and reduce childhood obesity rates. A question of comparable import, however, is whether healthier meals affect student achievement. Recent research demonstrates that the provision of subsidized school meals can significantly increase school test scores (Figlio and Winicki [2005]; Imberman and Kugler [2014]), but to date little evidence exists on how the quality of school meals affects student achievement. To answer the question of whether the quality of school meals affects student achievement we exploit longitudinal variation in California school district contracts with meal vendors to estimate differences-in-differences type specifications. Using five years of detailed test score data for California public schools, we find that contracting with "healthier" meal vendors significantly increases standardized test scores.

We combine two principal data sets from the California Department of Education, one covering school-level breakfast and lunch vendors and the other containing grade-level standardized test results. The vendor information covers the 2008 to 2013 academic years. For each California public school we observe whether the school had an outside contract with a meal provider for the school year, and if so the name of the provider and the type of contract. The vast majority of schools provide meals using "in house" staff, but a significant and growing fraction (approximately 10 percent) contract with outside vendors to provide meals. Crucial to our research design, there is significant turnover in vendors at the school level during our sample period.

We assign each vendor in our sample a "standard" or "healthy" vendor classification based on sample menus and nutritional information published by the vendors. Two different researchers independently classified the vendors as being either a low, medium, or high health vendor. A vendor received a healthy designation if each researcher classified the vendor as high, or if one researcher classified the vendor as high and one as medium. Otherwise the vendor was classified as standard. There is a high correlation between the two independent classifications, and the main results of the paper are robust to using either of the two classifications independently.

Estimating our main specification we find that contracting with a healthy meal vendor increases test scores by 0.03 standard deviations. This result is statistically significant and robust to the inclusion or exclusion of our time-varying covariates. In contrast there is no economic or statistically significant effect of contracting with a standard vendor.

We also estimate the effect of having a school lunch vendor separately for economically disadvantaged and non-disadvantaged students. We find that contracting with a healthy vendor has almost twice as large an effect on economically disadvantaged students (0.05 standard deviations) as it does on non-disadvantaged students. This may occur because economically disadvantaged students are more likely to eat school-provided meals. Economically disadvantaged students qualify for the National School Lunch Program (NSLP) which provide free or reduced rate school lunches. These students may also benefit more academically because economically disadvantaged students have poorer counterfactual nutrition intake.

To test whether differential trends in test scores between schools that contract with healthy vendors and other schools could drive our results, we conduct a series of tests in which we code a "placebo" treatment that activates one year prior to the year in which a school actually contracts with a healthy vendor. We find that the coefficient on the placebo treatment is close to zero (less than 0.01 standard deviations) and statistically insignificant. This implies that test scores are not trending differently prior to the year of the contract for schools that contract with healthy vendors. We also conduct a series of falsification tests in which we use our time-varying covariates as dependent variables. We find that changes in observable characteristics of schools are uncorrelated with new vendor contracts.

Although our estimated effects of 0.03 to 0.05 standard deviations are not large on an absolute scale, they potentially represent an exceptionally high benefit-cost ratio for a human capital investment. We find that it costs about $260 per student per year to contract with a healthy vendor relative to a standard vendor. Thus, overall it costs about $87 to increase a student's test score 0.01 standard deviations and about $52 to increase an economically disadvantaged student's test score 0.01 standard deviations. In comparison, experimental studies of class-size reductions and preschool programs have found that $1,000 of annual spending increases one students test scores by 0.01 to 0.02 standard deviations.

2 Data

The data for this project come from the State of California Department of Education. We use information on school-level breakfast and lunch vendors, and grade-level standardized test results. Each type of information is described in detail below.

2.1 Vendor Data

The breakfast and lunch vendor information are provided by the California Department of Education for the school years 2008-2009 to 2013-2014.[1] By law all food vendor contracts with public (K-12) schools in California must be approved by the CA Department of Education. The CA Department of Education retains information on all vendor contracts beginning with the 2008-2009 school year.[2] For each public school (including charter schools) in CA we know whether the school had an outside contract with a meal provider for the school year, the name of the provider, and the type of contract. The vast majority of vendor contracts are signed in the summer and cover the entire school year. A very small number of contracts cover less than the complete school year. These contracts correspond to the calender year and thus cover only a fraction (August-December or January-June) of the school year. We label these contracts "half-year" contracts. The estimation results are insensitive to the inclusion of schools with half-year contracts in our sample.

[1] The data were received as part of an official information request. We thank Rochelle Crossen for her assistance in facilitating the request and in interpreting the data.

[2] Contract information for school years prior to 2008-2009 were not retained when the CA Department of Education switched computer database systems.

There are four main types of school vendor contracts: "Vendor", "Food Service Management Company" (FSMC), "Food Service Consulting Company" (FSCC), and "School Food Authority". A "Vendor" contract is where a school contracts with an outside company that provides the meals that will be served at the school. However, school employees (i.e. cafeteria staff) still handle and serve the food including any additional prepping and cooking. In a FSMC contract, a private company prepares the meals and assists in staffing the school with cafeteria workers who serve the meals. In a FSCC contract, a private company provides "consulting services" on meal preparation and staffing, but does not provide any personnel for the jobs. Finally, a "School Food Authority" contract is when a school has a vending contract from another public school district. We do not distinguish between these four types of vending contracts in the main analysis of the paper. We refer to all four types of contracts as vending contracts.[3]

Detailed vendor contract information is available for a subset of the contracts. Contract details include meals provided (i.e. lunch, or breakfast and lunch), the dollar value of the contract, the number of other contract bidders (if any), the names of the companies which bid for the contract and were not selected, the dollar value of losing contracts, and the method which the contract bids are solicited (i.e. sealed bid or negotiation).[4] We use the contract

[3]Note that most private food companies in our sample have each type (Vendor, FSMC, and FSCC) of contract in our panel. That is, private food companies don't appear to specialize in a particular type of contract. Instead, the different contracts specify the exact services that the private company provides to the public school.

[4]The contract details are not available for all contracts for two reasons. First, school districts are only required to provide contract details to the state for the first year of a new contract. A contract can be renewed up to 4 times without having to issue a new contract. Second, school officials enter the contract information via a software program that

bid information to construct counterfactual estimates for the cost to improve state test scores by switching from a standard lunch provider to a healthy lunch provider.

We assign each vendor in our sample a "standard" or "healthy" vendor classification based on sample menus and nutritional information published by the vendors. Two different researchers independently classified the vendors as being either a low, medium, or high health vendor. A vendor received a healthy designation if each researcher classified the vendor as high, or if one researcher classified the vendor as high and one as medium. Otherwise the vendor was classified as standard. There is a high correlation between the two independent classifications and the main results of the paper are robust to using either of the two classifications independently.[5] Misclassification of a vendor as unhealthy when in fact it is a healthy vendor (or vice versa) is likely to reduce any estimated difference between the correlation of STAR scores and vendor type.

2.2 Number of School Lunches Served

National School Lunch Program (NSLP) data was obtained from the California Department of Education's Nutrition Services Division for the school

electronically stores the data in the CA Department of Education database. In practice, many of the data fields are missing for most of the new contracts. This is because, until recently, the CA Department of Education didn't have the staff to review the contract price and bid data entered into the system.

[5]Please refer to the Appendix for a more detailed discussion of the classification system, including links to sample menus and nutritional information. The Appendix also discusses several unsuccessful attempts to utilize 3rd party health ratings and/or health criteria. Surprisingly, we have (so far) been unable to find any established "healthy lunch" criteria we could uniformly apply to vendors, or any 3rd party health ratings for the private lunch providers.

years 2008-2009 through 2012-2013. The data reports the average number of NSLP lunches served per operating day in each school district. Averages are calculated monthly, so in order to obtain an annual measure for the average number of lunches served per day, we multiply the monthly averages by the number of operating days in each month and add up the monthly totals. The months of June and July are excluded from the total since these months may correspond to summer lunch programs that are managed separately. The annual total is divided by the total number of operating days in the year, again excluding June and July, to calculate an annual average of lunches served per day. Lastly, we divide the number of lunches served per day by the total enrollment in the school district to eliminate changes in lunches served due only to changes in the number of enrolled students.[6] Because we are interested in separately estimating the effect on disadvantaged students, we calculate averages for both total lunches served and free lunches. A student is eligible for a free school lunch if their family's income is less than 130% of the poverty level, and a reduced price lunch if their family's income is between 130% and 185% of the poverty level. The CA Department of Education refers to these students as "economically disadvantaged".

2.3 Academic Test Data

We use California's Standardized Testing and Reporting (STAR) test data. The STAR test is administered to all students in grades 2-11 each spring towards the end of the academic year. The publicly available test scores are

[6] Enrollment data by school is available from the California Department of Education. This data is aggregated to the school district level for our purposes.

aggregated by school and grade level.[7] The STAR test includes four core subject area tests: English/Language Arts, Mathematics, History/Social Sciences, and Science. In addition to the four core areas, there are a set of end-of-course (EOC) examinations (e.g. Algebra II, Biology).[8] The empirical analysis of the paper focuses on how the composite STAR score varies based on whether the school contracts with an outside vendor and on the healthiness of the food provided by the outside vendor. We use the standard deviation of each test (which differs by grade and year of test) as reported by California Department of Education to standardize each year's test scores. We then calculate the average test score (by year and grade) across all of the STAR test results taken by students in a particular grade in each school for each year.[9]

Average test scores are also available separately for students who qualify for reduced price and/or free school lunch under the NSLP. Students eligible for the reduced price or free lunches are the students most likely to eat the lunch offered at the school. Thus, we hypothesize that the academic benefit of having healthier school lunches would be largest for these students. Finally, school-level demographic and socioeconomic information is available

[7]We downloaded STAR test results for the years 1998-2013 from the California Department of Education website: http://star.cde.ca.gov/. Note that beginning with the 2013-14 school year that STAR testing was replaced with the California Assessment of Student Performance and Progress test.

[8]The subject areas tested vary by grade. English/Language Arts is tested in all grade years. Mathematics is tested in grades 2-7. History/Social Science is tested in grades 8 and 11. Science is tested in grades 5, 8, and 10.

[9]The qualitative results are robust to using only core test results, or in using just the English/Language Arts exam (which is the only exam taken by students in each grade). However, the point estimates are consistently lower in specifications that only use test results from the English/Language Arts exam (Appendix Table 7). This is consistent with other recent studies that separately measure the effect of access to school breakfast on test scores (e.g. Dotter [2013]; Imberman and Kugler [2014]).

for test-takers including: race, parental education, and students with English as a second language. We use this information to control for time-varying differences within schools in our main econometric model.

3 Empirical Specification

Our main empirical specification is a (five year) panel regression model.

$$y_{gst} = \beta_0 + \delta_H Healthy_{gst} + \delta_S Standard_{gst} + X_{st}\beta + \lambda_{gs} + \gamma_t + \epsilon_{gst} \qquad (1)$$

The dependent variable y_{gst} is the mean STAR test score across all tests for grade g in school s in year t. Before taking the mean across tests we divide each test score by the standard deviation of the test. Thus, the dependent variable is measured in STAR test standard deviation units.

Our independent variables of interest are whether a student test-taker is exposed to a standard or healthy outside lunch provider. The variable $Healthy_{gst}$ equals one if school s contracts with a healthy outside lunch provider in year t (and is zero otherwise). The variable $Standard_{gst}$ equals one if school s contracts with a standard outside lunch provider in year t (and zero otherwise). When both $Standard_{gst}$ and $Healthy_{gst}$ equal zero then the school does not contract with an outside lunch provider; the school's employees (i.e. cafeteria workers) prepare and serve the lunches.

The model includes school-by-grade (λ_{gs}) and year (γ_t) fixed effects. The school-by-grade fixed effects control for any characteristics in a given grade and school that are stable throughout the five year estimation period (e.g. school catchment area characteristics, school infrastructure, STAR test differences by grade, or school staffing levels and leadership). Year fixed effects control for common state-wide factors such as state economic conditions and differences in the STAR test that vary by year throughout the panel. Most specifications of the model also include X_{st}, a vector of school-level control variables that varies over time. These control variables include: the racial composition of test-takers at the school, the educational attainment of the parents of test-takers, and the proportion of economically disadvantaged students.

We estimate Equation (1) with standard errors robust to heteroskedasticity and clustered at the school district level. We cluster standard errors at the school district level since it is frequently the case that school lunch contracts are signed the same year for multiple schools in the same school district, and because school district officials often have to approve any school lunch contract in the district. Finally, our preferred specification uses the number of test-takers for each grade-school-year observation as weights in the regression. Using the number of test-takers allows us to recover the person-specific correlation between the type of school lunch served and academic performance as measured by the STAR test.[10]

The identifying assumption is, after controlling for time invariant school-by-grade factors, common state factors, and the vector of time varying school-level characteristics, that a school's decision to contract with an outside ven-

[10]Using the number of test-takers as weights will also provide an exact correction for heteroskedasticity if the variance of the error term for the underlying model that uses the test-taker as the unit of observation does not vary by test-taker.

dor for school lunch provision is uncorrelated with other school-specific, time-varying factors that affect student test performance. If this is true, then we can interpret the estimate for δ_S (δ_H) as the causal effect of contracting with a standard (healthy) school lunch provider on student learning, as measured by performance on the STAR test.

4 Results

4.1 Vendor Choice and Test-Taker Characteristics

Table 1 shows mean test-taker socioeconomic and racial characteristics for schools in two different samples: the *All School* sample and the *Contract School* sample. The All School sample includes all public CA elementary, middle, and high schools that report STAR scores. The Contract sample is limited to the subset of schools that have a school lunch vendor contract for at least one year in our 5 year panel. The means for each test-taker characteristic are calculated by first taking the 10 year (2004-2013) school-level mean.[11] In the All School sample the average school mean is then calculated separately for schools that do (Column 1) and do not (Column 2) contract with a vendor during our panel (2009-2013). Column (3) calculates the difference in means and provides the probability value (in parentheses) from a test of the null hypothesis that the means of the two groups are equal. The means are statistically different from each other at the 5% level for 9 of the 12 characteristics. For example, schools

[11] Note that STAR test-taker data are available for years before the start of our estimating panels. The length of our panel is determined by the availability of the vendor contract data.

that contract with a vendor during our sample tend to have fewer economically disadvantaged students, more gifted and talented students, a higher proportion of Asian students, and a higher fraction of students with parents who have a college degree. The results are similar for the Contract Sample.

Table 1 shows that student test-takers at schools that contract with an outside vendor are somewhat different than students at schools that don't contract with a vendor. Among schools that ever contract with a vendor, those schools that contract with a healthy vendor have different test-taker characteristics (on average) than those schools that contract with a standard vendor. These differences in test-taker characteristics in the two samples effects the generalizability of any association between test scores and vendor quality. The effect we measure could differ because the test-taking population among schools that don't contract with a vendor (or certain type of vendor) is different. Nevertheless, the differences in average characteristics between test-takers does not violate the identification assumption of Equation 1.

Table 2 shows how *changes* in the test-taker characteristics correlate with the timing of a vendor contract. We cannot interpret an observed correlation between vendor adoption and test score changes as a causal effect if changes in test-taker characteristics at a school can predict when a school contracts with an outside vendor. Table 2 displays the coefficient estimates from 14 different regressions using a version of Equation (1). In each of the first six columns we use a different test-taker characteristic as the dependent variable in place of test scores. In the last column we use the fitted values from a regression of test scores on all six test-taker characteristics (and year fixed effects) as the dependent variable. These fitted values summarize all of the test-taker characteristics, weighting each characteristic in relation to its correlation with test scores. All regressions in Table 2 include school-by-grade fixed effects and thus test whether within school-by-grade changes in student characteristics correlate with the time at which a school adopts an outside lunch provider.

Panel A of Table 2 estimates models using the All School sample, while Panel B uses the Contract School sample. None of the estimated coefficients are statistically significant at any level. The point estimates are small in magnitude, precisely estimated, and on net suggest negative selection into healthy vendors. The estimates in the last column reveal that adoption of a healthy vendor correlates with a statistically insignificant 0.01 standard deviation decline in predicted test scores, and the confidence intervals strongly reject any increase in predicted test scores above 0.01 standard deviations. The estimates for adoption of a standard vendor are also small and statistically insignificant. We interpret these results as initial evidence that changes in test-taker characteristics are uncorrelated with the timing of when a school contracts with a lunch provider.[12] Section 5 considers several additional tests of the validity of our identifying assumption.

[12] Appendix Table 7 shows estimates from a model that estimates the correlation between whether a school has a vendor using all of the covariates in a single regression. This allows us to jointly test the hypothesis that all characteristics are uncorrelated with the timing of vendor adoption. None of the coefficients on the characteristics are individually or jointly significant.

4.2 Vendor Choice and Test Scores

Table 3 shows estimation results for the effect of vendor quality on STAR scores. The first three columns estimate versions of Equation (1) on the Contract School sample, while the last three columns use the All School sample. Column (1) estimates the effect of contracting with a standard or healthy lunch vendor on test scores and includes only year and school fixed effects as controls. Column (2) adds school-by-grade fixed effects, while column (3) adds the vector of student test-taker characteristics. The point estimate for having a healthy vender on test scores, relative to no outside vendor, ranges from 0.030 to 0.034 standard deviations and is statistically significant at the 1% level in each of the three specifications. The estimate for a standard vendor is positive, but not statistically different from zero in each specification. The estimates for a healthy vendor from the All School sample are also statistically significant at the 1% level, consistent across specifications, and very similar to those estimated for the Contract School sample (ranging from 0.029 to 0.031). The estimates for the standard vendor are again positive, but not statistically significant.

The signs of the estimated coefficients for the student test-taker characteristics (in columns (3) and (6)) are what we would expect from previous literature. The omitted racial group variable is the percent of students at the school who are non-Asian minority. Relative to this group average test scores are higher at a school with a larger fraction of white and Asian students, but the fraction of Hispanic students has no statistically significant relationship with test scores. The coefficient magnitudes imply that a 10 percentage point increase in the share white (Asian) is associated with a 0.01 (0.05) standard deviation increase in average test scores. The omitted parental education group is the percentage of students at the school whose parents graduated from high school (but not from college). Relative to this group, test scores are lower at schools with a larger fraction of students whose parents who did not graduate from high school and higher at schools with a larger fraction of students whose parents graduated from college. Finally, average test scores are lower at schools that have a larger fraction of students from families near to or below the poverty level.[13]

The fact that we observe very similar point estimates for the vendor coefficients in columns (2) and (3) (and columns (5) and (6)) is consistent with the conclusion from Table 2. If student characteristics are important in predicting when a school contracts with an outside vendor then the coefficients in Table 2 should be statistically significant and the vendor estimates in Table 3 would differ between specifications with and with out these variables.

Table 4 investigates whether the effect of contracting with a lunch provider on STAR scores is different for economically disadvantaged and economically advantaged students. Recall that economically disadvantaged students are defined by the CA Department of Education as those students who qualify for reduced and/or free school lunch under the National School Lunch Program (NSLP) based on family income. We expect that disadvantaged students would be more likely to eat school lunch than their classmates who do not qualify for

[13] Recall that "Disadvantaged" implies that a student qualifies for the reduced or free school lunch program. Students qualify if their family income is less than 185% of the poverty level.

reduced and/or free school lunch. Thus, we hypothesize that the effect on test scores of eating a healthy school lunch should be greater for these students than for students who do not qualify for reduced and/or free school lunch. Table 4 shows evidence in support of this hypothesis.

Table 4 considers the Vendor Contract sample, but limits the sample to those schools which report average STAR scores for both economically advantaged and economically disadvantaged students.[14] Column (1) of Table 4 estimates the effect of contracting with a lunch vendor on the average test score for economically disadvantaged students. Column (2) estimates the effect on the average test scores for economically advantaged students, while column (3) estimates the effect for all students. The point estimate for contracting with a healthy vendor is 0.051 standard deviations for disadvantaged students and more than 50% larger than the estimate for advantaged students (0.033 standard deviations). Further, the healthy lunch vendor coefficient estimate for advantaged students lies outside (and below) the 95% confidence interval for the healthy vendor coefficient for disadvantaged students. The estimate for a healthy vendor for all students in column (3) lies in between those for disadvantaged and advantaged students.[15]

There is some evidence for a positive and statistically significant effect on test scores for disadvantaged students at schools that contract with a standard

[14]Due to privacy restrictions, the CA Department of Education only releases the average test score (for a school-grade-year) if there are at least 10 students of the particular socioeconomic group who take the test. There is a 25% reduction in the size of the sample due to these sample restrictions.

[15]The estimation results in the table are similar regardless of whether we include the test-taker control variables, whether we limit the sample to the common sample with both advantaged and disadvantaged scores, or if we use the All CA school sample.

lunch provider. This is suggestive that there could be a positive effect on test scores based purely on an increased caloric intake by disadvantaged students. For example, the outside vendor might do a better job of preparing the same meals (e.g. pizza) relative to the cafeteria workers, thereby making the school lunches more palatable to the students and leading to increased consumption. An alternative explanation is that even the unhealthy vendors are marginally more healthy than the average meal prepared by the cafeteria workers, and that this difference is statistically significant when we focus on the students most likely to eat these meals. We attempt to tease apart these possible explanations in Section 5 by looking at the number of lunches sold and school attendance.

4.3 Robustness Checks

Table 2 showed initial evidence that changes in test-taker characteristics are largely uncorrelated with the timing of when a school contracts with a lunch provider. In this section we further test the validity of our identifying assumption that a school's decision to contract with an outside vendor for school lunch provision is uncorrelated with other school-specific, time-varying factors that affect student test performance.

Equation (2) is a model that tests whether there is a correlation between test scores and contracting with a vendor in years before the vendor contract begins and in years after the vendor contract ends.

$$y_{gst} = \beta_0 + \sum_{\tau=-4}^{4} \delta_H^\tau Healthy_{gst}^\tau + \sum_{\tau=-4}^{4} \delta_S^\tau Standard_{gst}^\tau + X_{st}\beta + \lambda_{gs} + \gamma_t + \epsilon_{gst} \quad (2)$$

Equation (2) is identical to our main estimating equation except that we replace the single indicator variables for the year that a school contracts with a vendor ($Healthy_{gst}$ and $Standard_{gst}$) with a set of indicators ($Healthy_{gst}^\tau$ and $Standard_{gst}^\tau$) that also include indicators for the years before ($\tau < 0$) and after ($\tau > 0$) a school contracts with a vendor.[16] The indicator variables for a year before a contract are normalized to zero when we estimate Equation (2). Thus, the estimated coefficients δ_H^τ and δ_S^τ are interpreted as the change in test scores for students in grade g, school s, and year t relative to the year before a contract.

Figure 1 plots the estimated healthy (circles) and standard (squares) vendor event time coefficients and the 95% confidence intervals for the Contract School sample. The x-axis measures event time years (i.e. τ) and the y-axis measures test scores for all test takers. In a healthy vendor contract year there is an increase in test scores of 0.039 standard deviations relative to the year before a contract.[17] There is no evidence that increases in test scores precede contracting with a vendor, nor is there evidence for an upward pre-trend in test scores. The coefficient estimates for 2-4 years before a contract are close to zero and not statistically different from zero. Similarly, none of

[16]For example, $\tau = -4$ equals 1 if a school contracts with a vendor 4 years later (zero otherwise), and $\tau = 4$ if a school contracts with a vendor 4 years ago (zero otherwise).

[17]As a comparison, the estimate on test scores for the year of a healthy vendor contract from Equation (1) on the same sample is 0.034 (Table 3 column 3).

the estimated coefficients in the years after a contract ends are statistically significant.[18] Finally, none of the standard vendor coefficients are statistically significant.

There are two qualifications to the analysis in Figure 1. First, the event study coefficients towards the ends of our panel can be imprecisely estimated as there are fewer observations to identify these coefficients.[19] We address this concern by also estimating a model that pools the event time coefficients.[20] Second, we don't know whether a school contracts with a vendor in the years before our five year panel begins. This could lead our estimates to be biased if there is persistence in the test score effect after the vendor contract ends. The reason for this is that the model would incorrectly attribute the lagged effect on test scores (from having a vendor before our panel begins) as due to a new vendor contract in our panel.[21] We conclude that this is unlikely to be a concern since there is no evidence that the effect on test scores persists after a contract ends.

Table 5 shows the estimation results of five additional specifications that further test our identifying assumption and the robustness of our main test score results. Column (1) estimates Equation (2) except that the event time

[18]We also fail to reject F-tests that all of the coefficients in the years before a contract are equal to zero, and that all of the coefficients in the years after a contract are equal to zero.

[19]For example, the indicator for four years before a vendor contract can only equal one if a school contracts with a vendor in the last year of our panel. By contrast, an indicator for one year after a vendor contract ends could equal one for four of the five years in our panel.

[20]See Table 5.

[21]For example, Gallagher [2014] examines the effect on the take-up of flood insurance after a community is flooded using a model similar to Equation (2). Gallagher [2014] shows that the estimate for flood insurance take-up in the year of a flood is about 20% lower if the model fails to control for the lagged effect of a flood that occurs before the panel.

indicators for the years before and after a vendor contract are pooled for statistical power (e.g. Sojourner et al. [Forthcoming]). That is, the indicator *pre-trend* equals one if any of the indicators for $\tau \in [-4, -2]$ equals one, and *post-trend* equals one if any of the indicators for $\tau \in [1, 4]$ equals one.[22] If we estimate that *pre-trend* < 0, then this would suggest that school test scores are increasing even before the introduction of a new lunch vendor. We do not find evidence for any trends before or after a school contracts with a vendor. The estimated coefficient for *pre-trend* is close to zero and not statistically significant.

Column (2) of Table 5 considers a placebo test in the spirit of Equation (2) where we incorrectly consider the year before a vendor contract as the year of a contract (e.g. Currie et al. [2010]). We define *healthy placebo* (*standard placebo*) as equal to one if the school contracts with a healthy (standard) vendor in the following year. The estimated coefficients for both vendor placebos are close to zero and not statistically different from zero after controlling for the actual vendor years. There is no evidence that test scores begin to rise in the year before a school contracts with a vendor.[23]

Column (3) of Table 5 considers the sub-sample of schools from the Contract Sample that only ever contract with a standard vendor. As we would expect given earlier results the point estimate is not statistically different from zero. Column (4) considers the sub-sample of schools that only ever contract with a healthy vendor. The point estimate of 0.017 standard deviations is

[22]Note that the indicator for the year before a new vendor contract is still normalized to zero.

[23]Note that the estimated placebo coefficients are also close to zero and not statistically significant in a specification that doesn't condition on the actual vendor years.

statistically significant, but somewhat smaller than the estimate on the larger sample that includes all Contract Schools.[24] Finally, Column (5) tests whether test scores increase more in the 2nd, 3rd, etc. consecutive year of having a healthy lunch vendor.[25] To conduct this test, we add an indicator variable to Equation (1) that equals one if a school has a healthy vendor and it is at least the 2nd consecutive year of having a healthy vendor. We limit the sample to schools that only have healthy vendors and exclude the lowest level grade from these schools (e.g. we exclude grade 9 tests from high schools with grades 9-12).[26] The estimate for having a vendor for the 2nd, 3rd etc. consecutive year (interaction variable) is close to zero and not statistically significant (although the confidence interval is large).

5 Discussion

5.1 Number of Lunches Sold

The national Healthy, Hunger-Free Kids Act was passed by congress in 2010. The law increases the minimum nutritional standards that school lunches must meet. For example, the number of mandated servings of fruits and vegetables were increased, while at the same time restrictions were placed on the number

[24] There are 165 schools that contract with both a healthy and standard vendor during our panel. These schools are excluded from the samples estimated in columns (3) and (4).

[25] This would be the case if there were a year to year compounding effect of having a healthy lunch provider such that the learning, as reflected in test scores in the first year, prepares the student to do better in the 2nd year (over and above the effect of having a vendor in the 2nd year).

[26] The rationale for excluding the lowest level grade from each school is to ensure that students who continuously go to the same school would have at least two years of exposure to the vendor lunches.

of servings of French fries (Fed [2012]). A major goal of the law is to improve the health of school age children via a reduction in obesity (USD [2013]).

The first provisions of the law became binding beginning with the 2012-2013 school year. One criticism of the law is that improving the health content of the lunches may have the unintended consequence of reducing the number of students eating school lunches. A decrease in the number of meals served to students eligible for reduced price or free lunches would be most concerning as these students are considered most at risk for undernourishment and are the target population under the National School Lunch Program.

5.2 Student Health

We use physical fitness information on students in grades 5, 7, and 9 to examine whether having an outside lunch vendor affected obesity rates. The Physical Fitness Test (PFT), also called FitnessGram®, is given to students in grades 5, 7, and 9 each Spring in California. The PFT measures each student's body composition using the Body Mass Index (BMI).[27] The BMI is calculated using skin-fold calipers and measured in two locations (calf and triceps). The data are aggregated by school and grade level and indicate the percentage of students that have a BMI in the healthy fitness zone. The BMI healthy fitness zone differs for boys and girls and by age. For example, 15 year old boys are considered healthy if they have less than 25% BMI, while 15 year old girls are

[27] We downloaded 10 years (2004-2013) of PFT data directly from CA Department of Education website: http://www.cde.ca.gov/ta/tg/pf/pftresults.asp. Overall the PFT covers six fitness areas. We only use the BMI data. The other five areas are: Aerobic Capacity (measured by a one-mile run), Abdominal Strength (measured by "curl-ups", i.e. sit-ups), Upper Body Strength (measured by push-ups, or modified pull-ups), Back Strength (measured by trunk lift), and Flexibility (measured by sit and reach).

healthy if their BMI is less than 32%. Following Currie et al. [2010] we define overweight as the percentage of students falling outside the healthy zone.

We do not find any evidence that contracting with a healthy lunch provider reduces obesity among 5th, 7th, and 9th grade students. We estimate Equation (1) except that we use as a dependent variable the percent of students who are outside the healthy fitness zone (which we label as "obese") and restrict the sample to only grades 5, 7, and 9 from our Contract sample.[28] On average, 31.7 percent of the students in our sample are obese. Column (2) of Table 6 shows the estimation results. The point estimate for contracting with a healthy vendor is small (-0.10) and not statistically significant. We can rule out a 1.5 percentage point decrease and a 1.3 percentage point increase (using the 95% confidence interval) in the change of the proportion of students who are obese in years that the school contracts with a healthy vendor. The point estimate for an unhealthy vendor is also statistically insignificant with a similar confidence interval.

5.3 Heterogeneity by Grade

In this subsection we consider how the effect of having an outside lunch vendor varies by grade. For example, a school district may wish to spend more to contract with a healthy lunch vendor (rather than a standard vendor) if the district is particularly concerned about student performance on standardized tests in certain grades. Columns (3)-(5) of Table 6 estimate Equation (1) separately for students in elementary school (grades 2-5), middle school (grades

[28] The sample includes 4,006 grade-year observations at 910 schools.

6-8), and high school (grades 9-11). We separately pool students in all elementary, middle, and high school grades together, because any change in school lunch would be very likely to effect students in all grades in a school.[29] The majority of schools in California that contract with a vendor are elementary schools. As such, the elementary school estimates in column (3) are the most precise. There is a statistically significant 0.026 increase in test scores in years when a school contracts with a healthy vendor. The middle school estimate for a healthy vendor is of a similar magnitude, but not statistically significant. The largest estimated effect for a healthy vendor is for high school students. The estimates for the effect of a standard vendor are not statistically significant at the standard 5% level for any type of school.[30]

5.4 Policy Counterfactual

Public school administrators interested in improving the level of student learning and increasing test scores face a decision of how best to budget limited school resources. There are many potential changes in school policy that could improve learning. For example, school administrators could hire more teachers to decrease average classroom size (e.g. Krueger [1999]), lengthen the school day, increase teacher training (e.g. Angrist and Lavy [2001]), give bonus pay to teachers based on student test scores (e.g. Fry), or increase student ac-

[29]Note that our school definitions (e.g. middle schools include grades 6-8) are used as convenient labels and are not meant to imply, for example, that no middle school in our sample includes 5th grade students.

[30]The point estimate for a standard vendor is marginally significant at the 10% level for high schools (0.099 t-statistic). The large point estimate and marginal significance is suggestive that a standard vendor has a potential effect on high school test scores, but we are cautious in our interpretation of this coefficient given the small sample size and that the other specifications found no effect for the standard vendor.

cess to free or reduced price breakfast and lunch (e.g. Imberman and Kugler [2014]).[31]

Policies that have directed resources towards teachers have been shown to have a relatively large impact on student test scores in certain settings. The Tennessee STAR experiment, which reduced average class size for primary school students by one-third led to a 0.22 standard deviation test score increase, is a frequently cited benchmark (Krueger [1999]). Nevertheless, policies that direct resources towards teachers are often expensive and can be controversial (e.g. incentive pay). The Tennessee STAR experiment cost approximately $25 million (2013 $) with an implied cost of $3,009 (2013 $) per student placed in a smaller class.[32] Jacob and Rockoff [2011] highlight both the need and opportunity for cost-effective policies. Policies with relatively modest effects on student test scores may be more cost effective than policies with larger absolute effects.

We take advantage of contract-specific winner and loser bid information submitted to the CA Department of Education to calculate the cost differences between healthy and standard lunch providers. The average difference between healthy and standard lunch contracts is $260 (2013 $) per test-taker per school year. The CA school year is 180 school days. Thus, on average, a healthy school lunch contract costs about $1.44 more per student lunch than a lunch provided by a standard provider. Using this cost difference and an estimated effect of 0.034 standard deviations (Table 3 Column (3)) we find that it would

[31] This list highlights only a handful of policies and is not meant to be exhaustive.

[32] The original cost estimates reported by Krueger [1999] are adjusted to 2013 $ using the Consumer Price Index (CPI).

cost about $7,647 to raise a student's test score by one standard deviation by switching from a standard lunch provider to a healthy lunch provider. By way of comparison, it cost $13,678 to raise a student's test score by one standard deviation in the Tennessee STAR experiment.

6 References

Nutritional standards in the national school lunch and school breakfast programs, January 26 2012.

Usda proposes standards to provide healthy food options in schools, February 2013. URL http://www.fns.usda.gov/school-meals/healthy-hunger-free-kids-act.

Joshua Angrist and Victor Lavy. Does teacher training affect pupil learning? evidence from matched comparisons in jerusalem public schools. *Journal of Labor Economics*, 19(2):343–369, 2001.

Janet Currie, Stefano DellaVigna, Enrico Moretti, and Vikram Pathania. The effect of fast food restaurants on obesity and weight gain. *American Economic Journal: Economic Policy*, 2010.

Dallas Dotter. Breakfast at the desk: The impact of universal breakfast programs on academic performance. Working Paper, 2013.

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

Justin Gallagher. Learning about an infrequent event: Evidence from flood insurance takeup in the us. *American Economic Journal: Applied Economics*, 6: 206–233, 2014.

Scott A. Imberman and Adriana D. Kugler. The effect of providing breakfast on student performance. *Journal of Policy Analysis and Management*, 33:669–699, 2014.

Brian A. Jacob and Jonah E. Rockoff. Organizing schools to improve student achievement: Start times, grade configurations, and teacher assignments. *Brookings Institution*, September 2011.

Alan B. Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999.

Aaron J. Sojourner, Elton Mykerezi, and Kristine L. West. Evidence from adoptionsq-comp in minnesota. *Journal of Human Resources*, Forthcoming.

7 Tables

Table 1: Test-Taker Covariates for Schools that Contract with School Lunch Vendors

Sample:	All School Sample			Contract Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Vendor	No Vendor	Difference	Healthy	Standard	Difference
English	0.753	0.770	-0.017**	0.505	0.627	-0.122***
			(0.012)			(0.000)
Gifted and Talented	0.084	0.074	0.010***	0.038	0.084	-0.046***
			(0.000)			(0.000)
Disadvantaged	0.491	0.585	-0.094***	0.327	0.430	-0.103***
			(0.000)			(0.000)
Advantaged	0.507	0.413	0.094***	0.295	0.438	-0.143***
			(0.000)			(0.000)
Black	0.083	0.081	0.001	0.108	0.050	0.059***
			(0.754)			(0.000)
Asian	0.115	0.066	0.049***	0.098	0.105	-0.007
			(0.000)			(0.536)
Hispanic	0.422	0.471	-0.049***	0.186	0.398	-0.213***
			(0.000)			(0.000)
White	0.298	0.320	-0.022**	0.183	0.239	-0.056*
			(0.014)			(0.007)
Parent not HS grad	0.145	0.167	-0.022***	0.064	0.140	-0.076***
			(0.000)			(0.000)
Parent HS grad	0.199	0.206	-0.006*	0.120	0.182	-0.062***
			(0.064)			(0.000)
Parent college grad	0.189	0.145	0.044***	0.106	0.164	-0.058***
			(0.000)			(0.000)
ELD	0.027	0.029	-0.002	0.009	0.024	-0.015***
			(0.234)			(0.001)
Schools	1,016	9,198		385	798	

Table 2: The Correlation Between Test-Taker Covariates and the Timing of School Lunch Vendor Contracts

Dependent Variable:	(1) Disadvantaged	(2) White	(3) Asian	(4) Hispanic	(5) No HS	(6) College
Panel A. All school Sample						
Healthy Vendor	-0.029	-0.031	0.076	-0.125*	-0.045	0.029
	(0.074)	(0.062)	(0.077)	(0.081)	(0.024)	(0.018)
Standard Vendor	-0.152***	0.023	0.030	-0.079	-0.030	0.061***
	(0.055)	(0.039)	(0.022)	(0.054)	(0.021)	(0.022)
School and Grade Fixed Effects	No	No	No	No	No	No
Healthy Vendor	0.006	0.000	-0.002	0.002*	0.000	-0.003
	(0.009)	(0.004)	(0.002)	(0.001)	(0.004)	(0.002)
Standard Vendor	-0.005	-0.007	0.000	0.002	-0.001	-0.005
	(0.009)	(0.007)	(0.002)	(0.005)	(0.005)	(0.003)
School and Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. Contract School Sample						
Healthy Vendor	0.028	-0.030	0.039	-0.052	-0.008	0.001
	(0.047)	(0.038)	(0.047)	(0.058)	(0.019)	(0.013)
Standard Vendor	-0.096	0.026	-0.009	-0.003	0.007	0.033
	(0.058)	(0.049)	(0.041)	(0.059)	(0.020)	(0.021)
School and Grade Fixed Effects	No	No	No	No	No	No
Healthy Vendor	0.007	-0.001	-0.003*	0.005	0.001	-0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Standard Vendor	0.000	-0.010	0.001	0.006	-0.001	-0.005
	(0.009)	(0.008)	(0.003)	(0.005)	(0.005)	(0.004)
School and Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Table 3: The Effect of Vendor Choice on Standardized Test Scores

Dependent Variable:	Standardized Test Score					
Sample:	Contract Schools			All Schools		
	(1)	(2)	(3)	(4)	(5)	(6)
Healthy Vendor	0.030***	0.031***	0.034***	0.029***	0.029***	0.031***
	(0.009)	(0.010)	(0.017)	(0.007)	(0.008)	(0.008)
Standard Vendor	0.016	0.014	0.017	0.002	0.001	0.002
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)
White			0.231**			0.232***
			(0.092)			(0.035)
Asian			0.623***			0.625***
			(0.136)			(0.062)
Flag Asian			0.006			0.005
			(0.011)			(0.004)
Hispanic			-0.028			-0.130*
			(0.149)			(0.051)
No High School			-0.201***			-0.168***
			(0.065)			(0.026)
College			0.284***			0.208***
			(0.054)			(0.043)
Disadvantaged			-0.114*			-0.141***
			(0.060)			(0.027)
Grade Fixed Effects		X	X		X	X
Covariates Included			X			X
R-squared	0.703	0.871	0.874	0.711	0.875	0.878
Obs	15,382	15,382	15,382	136,090	136,090	136,090
Schools	1,028	1,028	1,028	9,324	9,324	9,324

Table 4: The Effect of Vendor Choice on Standardized Test Scores by Socioeconomic Status

Dependent Variable:	Star Score Standard Deviations		
	(1)	(2)	(3)
Sample:	Disadvantaged	Advantaged	All
Healthy Vendor	0.051***	0.033***	0.039***
	(0.007)	(0.010)	(0.009)
Standard Vendor	0.033*	0.014	0.020
	(0.019)	(0.017)	(0.016)
Grade Fixed Effects	X	X	X
Covariates Included	X	X	X
R-squared	0.810	0.833	0.864
Obs	11,492	11,492	11,492
Schools	827	827	827

Table 5: The Effect of Vendor Choice on Standardized Test Scores: Robustness Checks

Dependent Variable:	Standardized Test Score				
	(1)	(2)	(3)	(4)	(5)
Model:	Trends	Placebo	Standard Only	Healthy Only	Accumulative
Healthy Vendor	0.027***	0.032**		0.017**	0.018*
	(0.008)	(0.011)		(0.008)	(0.010)
Standard Vendor	0.016	0.016	0.011		
	(0.017)	(0.015)	(0.019)		
Pre-trend Healthy Vendor	-0.007				
	(0.018)				
Pre-trend Standard Vendor	0.011				
	(0.015)				
Post-trend Healthy Vendor	-0.016				
	(0.015)				
Post-trend Standard Vendor	0.011				
	(0.011)				
Healthy Vendor Placebo		0.005			
		(0.009)			
Standard Vendor Placebo		0.009			
		(0.013)			
Healthy Vendor 2+ Years					-0.011
					(0.027)
Grade Fixed Effects	X	X	X	X	X
Covariates Included	X	X	X	X	X
R-squared	0.874	0.874	0.880	0.871	0.875
Obs	15,382	15,382	9,845	3,292	2,572
Schools	1,028	1,028	637	226	219

Table 6: The Effect of Vendor Choice on the Number of Lunches Sold, Body Mass Index, and Test Scores by School Grade

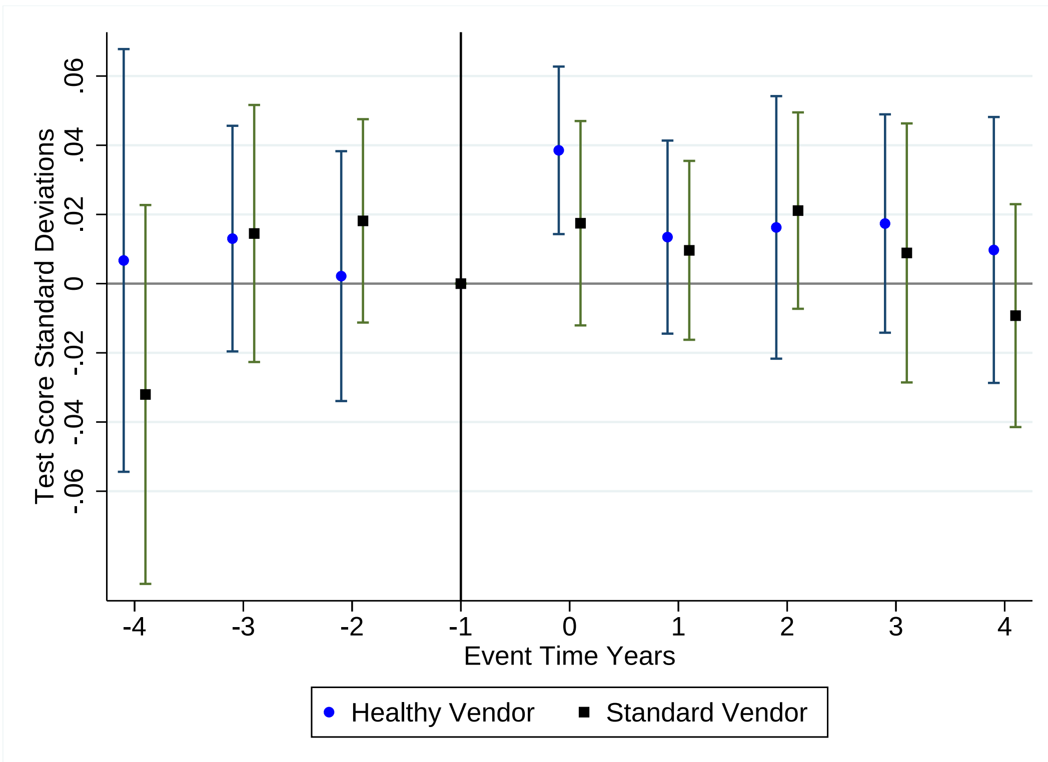
Model:	(1) Number of Lunches	(2) Body Mass Index	(3) Elementary School	(4) Middle School	(5) High School
Healthy Vendor		-0.104	0.026**	0.023	0.051***
		(0.705)	(0.011)	(0.017)	(0.013)
Standard Vendor		-0.267	-0.008	0.028	0.042*
		(0.805)	(0.019)	(0.024)	(0.025)
Grade Fixed Effects	X	X	X	X	X
Covariates Included	X	X	X	X	X
R-squared		0.816	0.857	0.883	0.841
Obs		3,934	9,496	3,737	2,149
Schools		902	665	486	230

Table 7: (Appendix) The Effect of Vendor Choice on Test Scores:
Alternative STAR Test Measures

Dependent Variable:	Core Tests			ELA Tests		
	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Disadvantaged	Advantaged	All	Disadvantaged	Advantaged	All
Healthy Vendor	0.032***	0.013	0.023***	0.031***	0.004	0.020**
	(0.008)	(0.013)	(0.009)	(0.007)	(0.013)	(0.009)
Standard Vendor	0.003	-0.020	-0.011	0.001	-0.023	-0.013
	(0.016)	(0.020)	(0.018)	(0.018)	(0.021)	(0.019)
Grade Fixed Effects	X	X	X	X	X	X
Covariates Included	X	X	X	X	X	X
R-squared	0.851	0.892	0.907	0.852	0.888	0.903
Obs	11,489	11,489	11,489	11,416	11,416	11,416
Schools	827	827	827	827	827	827

8 Figures

Figure 1: The Effect of Healthy and Standard Vendors on Test Scores



Notes: This figure depicts point estimates for treatment leads and lags with their corresponding 95% confidence intervals. Point estimates come from a weighted regression using number of test-takers per observation as weights. Regression includes year and school fixed effects with errors clustered at the school-district level. Sample is limited to the subset of schools that contract with any vendor at any point during our sample period.

9 Appendix

9.1 Heterogeneity and Robustness