Universal Access to Free School Meals and Student Achievement: Evidence from the Community Eligibility Provision

Krista Ruffini †

February 17, 2020

Abstract

The school meals program is the largest nutritional assistance program for school-aged children. Whereas program eligibility was historically determined by family income, recent reforms allow schools to offer free meals to all students. This paper evaluates the effect of the Community Eligibility Provision, the largest schoolwide free meals program, on academic performance. I leverage within- and across-state variation in the timing of CEP participation and find universal free meals increases breakfast and lunch participation by 38 and 12 percent, respectively. Math performance improves in districts with baseline low free meal eligibility, particularly among racial/ethnic groups with low income-based participation rates.

*I would like to thank Hilary Hoynes, Marianne Bitler, Avi Feller, Nora Gordon, Rucker Johnson, Enrico Moretti, Sean Reardon, Jesse Rothstein, May Lynn Tan, Christopher Walters, two anonymous reviewers, and participants at the GSPP research seminar, All-California Labor Conference, and Russell Sage Foundation for helpful comments and suggestions, and staff from the D.C., Georgia, Illinois, Kentucky, Maryland, New York, and West Virginia state departments of education for providing assistance with the state CEP and meal participation data. Miki Bairstow provided outstanding research assistance. Support from the Russell Sage Foundation and W.T. Grant Foundation (Grant #83-17-04) is gratefully acknowledged. All errors are my own.

†Goldman School of Public Policy, University of California, Berkeley. Email: k.j.ruffini@berkeley.edu

1 Introduction

The National School Lunch Program (NSLP) and School Breakfast Program (SBP) – jointly referred to as the school meals program – are the largest nutritional assistance programs serving school-aged children. On a typical school day, more than 30 percent of 5-17 year-olds receive free lunches, and more than 20 percent receive free breakfasts (USDA, 2018b).[1] These programs represent a large share of students' nutritional intake, as children consume up to half of their daily calories at school (Gleason and Suitor, 2001). School meals also provide a relatively generous income subsidy to low-income families: a student receiving free breakfasts and lunches pays approximately $4.50 a day ($800 per school year) less than students paying the full price.[2]

Despite the size and importance of the school meals program, evaluating the causal effect of school-based nutritional assistance on child outcomes presents empirical challenges. Until recently, there was little program variation across schools or over time: most schools implemented the lunch component within a short time period, the federal government establishes requirements for all schools, and nearly all schools participate. At the student level, family income determines payment rates. Accordingly, children who receive free meals are systematically more disadvantaged than ineligible students, which complicates comparisons of eligible and ineligible children. While the existing literature finds school meals increase food consumption and nutritional intake (Bhattacharya et al., 2006; Schanzenbach, 2009; Gleason and Suitor, 2001; Nord and Romig, 2006; Gundersen et al., 2012), results for other outcomes are more mixed (Dunifon and Kowaleski-Jones, 2003; Frisvold, 2015; Hinrichs, 2010; Meyers et al., 1989).

Recent reforms have transformed the school meals program from income-based assistance to more universal access by allowing schools and districts to offer free meals to all students,

[1] An additional five percent receive reduced-price meals at a deeply subsidized rate.

[2] In comparison, the average daily per-person SNAP benefit for families with children is about $4.00 (USDA, 2018a). The size of the school meals program is also apparent by examining federal outlays. In Fiscal Year 2017, the federal government allocated about $16.8 billion to school nutrition programs, compared to $15.0 billion in Title I funding and $28.0 billion in pro-rated SNAP benefits to children (USDA, 2017; US Department of Education, 2017; USDA, 2018a)

regardless of a student's family income. The shift towards school-based assistance has fundamentally altered the nature of the program: in the 2019 school year, more than a quarter of school-aged children attended a school offering universal meals, a marked increase from less than one percent in 2012 (Food Research and Action Center, 2019). This paper examines the most recent and largest such reform, the Community Eligibility Provision (CEP). Although universal programs formally increase access to free school meals, schools and districts with relatively high free meal participation under the traditional program (those that were *de facto* approaching universal provision) have the greatest incentives to participate. As CEP participation is voluntary, *ex-ante* it is unclear whether moving to universal assistance will affect meal participation or student test scores in these schools and districts.

In order to examine the effect of schoolwide free meals on consumption and student performance, I compare changes in early-adopting districts to those adopting later. This differences-in-differences approach relies on the fact that although CEP is a federal program, not all schools and districts became eligible for or adopted schoolwide free meals at the same time. Districts became eligible for CEP over a four-year period depending on state, and within states, participation among eligible schools and districts has increased over time. Importantly, this framework accounts for selection into CEP by limiting the analysis sample to "ever-adopting" districts – those with similar observed and unobserved participation incentives. If the timing of CEP adoption is uncorrelated with changes in potential performance, this approach identifies the causal effect of universal access to free meals.

To evaluate whether the timing of CEP adoption is plausibly exogenous among these ever-adopting districts, I conduct two complementary analyses. First, I explore whether baseline characteristics, such as district resources and economic well-being, systematically differ between early- and late-adopting districts. Here, I find that districts adopting in the first pilot year have slightly higher poverty rates and worse academic performance than areas adopted later, but economic conditions are not differently *trending* for the earliest adopters. Second, I present event study analyses illustrating trends in academic performance before and after CEP adoption. These plots show math performance was not systematically trending for black and Hispanic students prior to CEP adoption. On the other hand, this analysis suggests math performance among white students was improving prior to implementation. Therefore, while the timing of CEP adoption is more likely exogenous from the perspective of non-white students' performance trajectories, the findings for white students should be interpreted cautiously.

My findings are twofold. First, using administrative meal count data from six of the eleven pilot states, I find that even among districts with high baseline free meals eligibility, CEP increased the number of breakfasts and lunches served by approximately 38 and 12 percent, respectively.

Second, I examine how schoolwide free meals affects student performance in all CEP-participating districts nationwide. Improvements in math achievement vary by the share of students gaining access to free meals, with CEP improving performance in districts with the lowest eligibility rates under the traditional program, but not significantly affecting test scores in districts with high baseline free meal eligibility. These patterns are consistent with the nature of the program. Specifically, the full sample of participating districts includes areas with high free meal participation rates under the traditional program and these areas experienced little effective change in access under CEP.

In order to focus on districts that experienced the largest changes in access to free meals, I divide the sample of CEP-participating districts at the median baseline share of students qualifying for free and reduced meals – approximately 58 percent. For the "exposed" districts with relatively low baseline eligibility rates, CEP modestly improved math performance by about 0.02 standard deviations. Scaling these performance improvements by the share of students gaining access to free meals (32 percent) implies that access to free school meals improve math performance by approximate 0.05 standard deviations. Within the exposed subsample, improvements are concentrated among elementary and Hispanic students. The subgroup analyses are consistent with CEP providing benefits to students gaining access to free meals, as Hispanic students had relatively low free meal eligibility rates under the traditional program (Chaparro et al., 2014; Goerge et al., 2009). In contrast to modest improvements in math performance, changes in reading performance are more sensitive to the specification, smaller in magnitude, and generally statistically insignificant.

Math improvements follow an inverse U-shaped pattern over the "exposure" distribution. CEP districts with the lowest baseline eligibility rates tended to adopt CEP in some – but not all – schools, resulting in relatively small increases in free meal access at the district level. On the other end of the distribution, districts with high baseline eligibility rates experienced little change in access, as most students were already eligible for free meals. Districts in the middle of the distribution – those with baseline eligibility rates between approximately 50 and 60 percent – were able to expand free meals to a relatively large share of students. These districts are those that also tended to have the largest improvements in math performance.

There are several channels by which schoolwide free meals may affect student performance. First, students who are otherwise income-eligible, but who did not complete the required paperwork, gain access. Second, higher-income students become eligible free meals. Family resources increase for both of these students, which may yield academic benefits independent of any nutritional changes. Third, if universal access to free meals improves behaviors, it may reduce classroom distractions and increase teaching time, benefiting students whose nutritional consumption does not change. Fourth, since all students receive free meals under CEP, family income may become less salient, which could reduce stigma. Fifth, if CEP participation is determined by financial considerations – either from lower administrative costs or greater federal revenue – districts may provide additional educational supports. These channels are not mutually exclusive, and although the available data do not allow me to fully disentangle among possible mechanisms, results do not meaningfully change after accounting for changes in district resources, indicating the findings are not solely due to concurrent changes in financial or instructional resources.

This paper builds on a burgeoning literature that uses quasi-experimental variation to estimate the effects of nutritional assistance on health and economic outcomes. Much of the existing research examines family-based assistance through the Supplemental Nutrition Assistance Program (SNAP). For example, Hoynes et al. (2016) find access to SNAP in childhood improves adult outcomes, and Gassman-Pines and Bellows (2018) and Gennetian et al. (2016) find that greater SNAP resources improve short-term student performance and behaviors. The relationship between school-based assistance and children's outcomes is more mixed. While Schanzenbach (2009) finds school meals slightly increase obesity rates, Gleason and Suitor (2001), Schanzenbach and Zaki (2014), and Bhattacharya et al. (2006) find school meals improve nutritional intake. In the long-term, Hinrichs (2010) finds that greater exposure to school lunches increases educational attainment. Examining the short-term effects of these programs can help disentangle whether any long-term benefits arise directly through academic achievement, through latent health benefits, or non-cognitive improvements.

This analysis makes several contributions to the existing literature. First, this paper examines variation in access to free meals that is related to the characteristics of the surrounding area, but not driven by whether a particular student's family faces economic hardship. Second, the existing work on schoolwide free meal programs is largely limited to the first two years of implementation. As the first districts to adopt CEP are treated for four years in my sample, I am able to explore whether the marginal benefits of nutritional assistance increase or decrease with greater program experience.

In addition, previous work on universal meals almost exclusively focuses on a single, urban school district (Dotter, 2013; Imberman and Kugler, 2014; Schwartz and Rothbart, nd), or single state (Fuller and Comperatore, 2018; Kho, 2018; Gordanier et al., 2019; Davis and Musaddiq, 2018). This paper complements the existing research by examining how universal meals affect performance in both rural and urban districts for the near-universe of public school districts, and provides some of the first evidence on the extent to which the experiences of a single state or district reform may apply more broadly. Importantly, the national-level data and staggered adoption period permit a rich set of controls for other state-level changes occurring over the analysis period. To this point, I find improvements in math performance for exposed subgroups are robust to accounting for state-specific trends and leveraging only within-state variation in the timing of CEP adoption. I also find similar effects across geographic regions and states, and between rural and urban areas. In addition, I explore whether the findings are driven by other changes in the school environment by examining how CEP affected district resources and the types of students attending CEP districts and schools. Here, I show total per-pupil expenditures and federal non-nutrition district revenue did not significantly change following implementation. Changes in the student body composition – fewer Hispanic students and greater racial/ethnic segregation – account for no more than 10 percent of the observed improvements.

This paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 overviews the CEP reform, and outlines the channels by which school-based assistance can affect student performance. Section 4 describes the data and methodology. Section 5 presents results and Section 6 concludes.

2 Existing Literature on School Meals and Universal Provision

Food insecurity, defined as inadequate nutritional access, is associated with poor health and impaired social, emotional, and cognitive development (Howard, 2011). A growing body of research finds that nutritional assistance reduces children's food insecurity (Ratcliffe et al., 2011; Mabli and Worthington, 2014; Arteaga and Heflin, 2014; Bhattacharya et al., 2006; Gleason and Suitor, 2001; Frisvold, 2015; Fletcher and Frisvold, 2017; Gundersen et al., 2012). Even accounting for these programs, however, about 16 percent of families with children are food insecure, as household-based assistance often does not cover food costs (Hoynes et al., 2015; Coleman-Jensen et al., 2017) and not all income-eligible students participate in SNAP or the school meals program (Ganong and Liebman, 2013; Domina et al., 2018; Coleman-Jensen et al., 2017). Both prevalence of food insecurity and incomplete take-up suggests there is scope for schoolwide provision to improve children's health and school performance.

The traditional school meals program provides subsidized meals to lower-income school-aged children, with each student's required payment determined by family income: Children with family income below 130 percent of the federal poverty level pay $0 for school breakfasts and lunches, while children in families up to 185 percent of the federal poverty level pay no more than 40 cents. Higher-income students can purchase a meal at the "paid" rate, set by each district and averaging about $3.50 for middle school students. The federal government reimburses schools based on the number of free, reduced, and full-price meals served, with reimbursement rates shown in Table 1. Figure 1 shows that both the breakfast and lunch programs serve a large share of children, and participation in the free component has grown over time while participation in the paid component has remained relatively steady (breakfast) or declined (lunch).

The existing empirical literature has found mixed results on the effect of the traditional school meals program on student performance. For example, Dunifon and Kowaleski-Jones (2003) find free lunch participation does not significantly change student academic performance, while other work finds slight improvements following greater access to school breakfasts (Meyers et al., 1989; Frisvold, 2015) or more nutritious lunches (Anderson et al., 2017).

Before CEP, schoolwide free meal programs were largely district-initiated efforts that usually only provided free breakfasts. Many of these reforms also changed how meals were provided, for example, by serving breakfast during instructional time, rather than before school (e.g. "Breakfast in the Classroom"). A series of papers examines the effects of these early endeavors and finds universal, in-classroom breakfasts improve math and reading scores (Imberman and Kugler, 2014; Dotter, 2013). On the other hand, schoolwide free breakfast programs that maintain traditional serving methods increase participation, but do not improve performance (Bartlett et al., 2014; Schanzenbach and Zaki, 2014; Leos-Urbel et al., 2013).

With the available data, I am unable to determine whether CEP coincided with changes in how meals were offered. Surveys of school administrators suggest CEP may have affected both access to school meals and how these meals were offered. While most schools continued to operate a traditional "line/cafeteria" service under CEP, about one-third of districts expanded offerings to in-classroom and "grab-and-go" options (Logan et al., 2014). To the extent that CEP changed *how* meals are served, results in this paper should be interpreted as the "dual" treatment of universal provision and growing likelihood of adopting non-traditional serving methods.

Closely related to this paper, a number of studies examine the effect of schoolwide free breakfasts and lunches within a single state or district through CEP and other federal initiatives by comparing schools that opt to participate in schoolwide programs to those that do not (either conditional or unconditional on eligibility). Consistent with my results, the existing work tends to find universal free meals modestly improve math performance, particularly for elementary school students, with mixed results on reading and for middle school students (Gordanier et al., 2019; Kho, 2018; Fuller and Comperatore, 2018). Also consistent with benefits being concentrated among populations with low income-based eligibility rates, Schwartz and Rothbart (nd) find schoolwide lunches confer particularly large benefits for students who were income-ineligible ineligible under the traditional program. Finally, using a similar empirical approach to this paper, Gordon and Ruffini (nd) examine non-academic outcomes and find CEP reduced suspension rates among white, male elementary students. My results are consistent with CEP providing larger benefits for younger students, as well as those living in areas with greater unmet need.

The present study builds on the existing literature in three ways. First, it provides national-level estimates of schoolwide free meals by examining changes in district performance across the entire country. To the extent that state- and district-level evaluations reflect idiosyncratic local decisions, these national results are arguably more generalizable for policymakers contemplating program reforms. Second, by exploiting variation in the timing of adoption, rather than participation decisions, this paper relies on the identifying assumption that the timing of implementation, rather then whether to implement, is uncorrelated with potential gains. Finally, it broadens our understanding of which outcomes and student groups stand to benefit from schoolwide free meal programs by exploring heterogeneous treatment effects by student and area characteristics.

3 Community Eligibility Provision

3.1 CEP Program Details

The Community Eligibility Provision (CEP) is the largest schoolwide free meals program. In the 2019 school year, more than a quarter of school-aged children attended a CEP school (Food Research and Action Center, 2019).[3] CEP eligibility is based on a school or district's "identified student percentage" (ISP), the share of students who receive another form of income-based assistance, primarily SNAP.[4] Schools and districts with an ISP of at least 40 percent can choose to adopt CEP, and within a district, a subgroup of schools can "pool" ISP and elect to receive CEP as a "group."[5] Over the 2012-2015 period, approximately 60 percent of participating districts fully participated, and slightly more than half of students attended a CEP school in partially-participating districts.

Important for my identification strategy, districts became eligible to implement CEP at different times over a four-year window. The rollout order was based on state and determined by the Secretary of Agriculture to ensure "an adequate number and variety of schools and [districts] that could benefit from [CEP]". Districts in Illinois, Kentucky, and Michigan became eligible to participate in the 2012 school year; districts in the District of Columbia, New York, Ohio, and West Virginia were newly eligible in 2013; districts in Georgia, Florida, Maryland, and Massachusetts became eligible in 2014; and districts in the remaining states became eligible in 2015 (Figure 2).

Among eligible districts, about one-third participated by 2015, ranging from 0 percent

[3]Earlier schoolwide meal options include Provisions 1-3 which provide reimbursement according to base year shares of FRP students (USDA, 2002). These options are most beneficial to schools where nearly all students are income-eligible. For districts previously implementing Provisions 1-3, CEP did not change free meal access, but provided an alternative federal reimbursement. If schools aim to maximize federal revenue, Provision 1-3 schools that take-up CEP should experience a (weak) increase in federal revenue. These districts are included in my analyses only if they adopted CEP by 2017 and had a baseline FRP eligibility rate below 57.9 percent between 2009 and 2011 (e.g.: were not operating universal programs prior to 2011).

[4]Students receiving TANF or the Food Distribution Program on Indian Reservations, or who are foster youth, runaway youth, homeless, or migrants are also included in ISP.

[5]ISP is also referred to as "categorically-eligible" share or the fraction "directly certified". To see how schools may "pool" ISP, consider the following example: if one school in a district has an ISP of 20 percent and another (with equal enrollment) has an ISP of 60 percent, the two schools can combine ISP and be treated as a CEP participant with 40 percent ISP.

in New Hampshire to 81 percent in Montana (Neuberger et al., 2015). CEP participation has also increased within states over time. For example, in my sample, approximately 5 percent of districts in Kentucky, Illinois, and Michigan had at least one participating school in the first year of eligibility (2012). By the fourth year of eligibility in 2015, this figure had increased to 18 percent.

Both financial and student eligibility considerations affect a district's participation incentives. The federal government reimburses CEP participants at 1.6 times ISP, up to a maximum of 100 percent. For example, a district with an ISP of 40 percent receives federal reimbursement at the free meal price for 64 percent of the meals served. The remaining 36 percent are subsidized at the paid price. Since CEP schools forgo revenue from students who previously received paid meals, local sources cover any remaining costs, and these additional costs to districts reduce participation incentives. In contrast, areas with an ISP of at least 62.5 percent receive the full federal subsidy for all meals. Beyond 62.5 percent ISP, districts receive full federal reimbursement under CEP, but the financial benefit of CEP decreases since these districts were already receiving a high reimbursement rate under the original program.

In addition to potentially changing federal reimbursement, CEP increases the number of students with access to free meals. Schools and districts with the lowest baseline eligibility rates experience the largest increases in access. For example, a school with a 64 percent FRP share under the traditional formula would increase free meal access by 56 percent (36 percentage points) under CEP. On the other hand, a school with a 100 percent FRP share would see no change in access (regardless of ISP).

These incentives shaped participation decisions. During the pilot period, administrators in both participating and non-participating districts cited financial concerns or reimbursement rates as one of the three most important factors in deciding whether to participate, and approximately 80 percent stated that CEP would increase access to healthy foods (Logan et al., 2014). Empirically, Figure 3 shows that while districts with a higher baseline eligibility rate are most likely to participate in CEP, but even among the highest-poverty schools, only about 60 percent participate. More generally, as CEP districts have systematically higher baseline eligibility rates than those that do not participate (both those eligible and ineligible), participating districts are unlikely to be a random sample. In order to compare districts with similar observable and unobservable incentives to participate, my main specifications restrict the sample to districts with any school participating in CEP by 2017 and compare districts that adopted CEP relatively early to those that adopted later.

3.2 Conceptual Framework

There are several channels through which schoolwide free meals may affect average district academic performance. First, income-eligible students who do not complete enrollment paperwork gain access to the program. Second, higher-income students become eligible for free meals. For both of these groups, universal free meals increase family resources available for other food expenditures and consumption goods, which may benefit children.

Third, students' classroom experiences depend both on their own behavior and their peers' actions and classroom disruptions reduce the learning time of all students (Lazear, 2001). The literature shows that food insecurity is associated with worsened externalizing behaviors (Alaimo et al., 2001) and disruptive peers lead to worsened labor market outcomes for other students (Carrell et al., 2018). Therefore, if CEP lowers food insecurity, it may improve behaviors or reduce classroom distractions, increasing effective teaching time and benefiting students whose nutritional consumption does not change.

Fourth, since all students receive free meals in CEP schools, family income may become less salient, or consuming a school meal may become less stigmatizing, resulting in a more inclusive learning environment. Early focus groups suggest stigma reduced school meal consumption among income-eligible students (Glantz et al., 1994), and previous work examining the introduction of free meals in New York City finds increased participation among all students regardless of a student's initial FRP eligibility, consistent with universal meals reducing stigma (Leos-Urbel et al., 2013; Schwartz and Rothbart, nd).

Fifth, CEP may lower schools' administrative costs by reducing the need to track individual free meal eligibility and participation. Districts may reallocate these cost savings to resources that directly improve student performance.[6] Although I am unable to fully disentangle among these five channels with district-level data, results are very similar to the baseline specifications when controlling for district personnel resources and revenue, indicating the findings are not solely due to changes in financial resources.

4 Measuring CEP Participation and Achievement

4.1 CEP Participation

I combine information from several sources to estimate the effect of CEP on student performance.[7] I obtain CEP participation data for public and public charter schools from state educational agencies for the 2012 through 2014 school years and the USDA Food Research Action Center (FRAC) for the 2015 through 2017 school years. Within districts, there is some variation across grade levels in CEP adoption. Elementary schools have higher participation rates than middle or high schools: In districts with any CEP adoption, about 97 percent implemented CEP in at least one elementary school and about 80 percent implemented in at least one middle school.[8] In order to obtain a district measure of participation specific to each grade, I aggregate the yearly school-level participation information to the district-grade-year level.

[6] Although administrative costs may fall under CEP, the program's effect on net district revenue is ambiguous. There are two parameters shaping financial incentives, depending whether districts aim to maximize federal revenue or total nutritional assistance revenue (from students plus the federal government): ISP and FRP shares. First, districts with ISP rates below 62.5 percent receive less than 100 percent federal reimbursement and lose revenue from students who previously received school meals at the paid price. Second, districts with a FRP-ISP ratio above 1.6 receive less federal funds under CEP than the traditional program. Participation is expected to be lower for districts with either a FRP-ISP ratio above 1.6 or an ISP below 62.5 percent. Among such districts that do participate, the higher costs of the meals program may reduce funds available for other educational services. On the other hand, districts with an ISP of at least 62.5 percent are weakly better off under CEP.

[7] The data appendix describes each source in greater detail.

[8] Based on data from the state of Maryland, most districts with incomplete participation would financially benefit from additional CEP coverage through 2019, suggesting strategic applications are a negligible concern over the analysis period.

4.2 Student Performance

In order to obtain a measure of academic performance that is comparable across states and over time, I use a novel dataset from the Stanford Education Data Archive (SEDA). These data address several issues that have precluded sub-state comparisons of student achievement. In particular, data from the biennial National Assessment of Educational Progress (NAEP) does not include all schools and the universe of NAEP-tested schools changes each survey year. Both of these features limit comparsions of performance across districts or schools over time. Second, school-level proficiency data required of most states under the No Child Left Behind (NCLB) and the Every Student Succeeds Act (ESSA) are unreliable for cross-state comparisons, as each state designs its own test and proficiency metric, both of which substantially changed over the CEP implementation period.[9]

The SEDA data overcomes many of the limitations of the NAEP and NCLB data by using information from both sources. First, restricted-use, school-level NCLB proficiency data for the 2009 through 2015 school years are aggregated to the district-grade-year level. As detailed in Reardon et al. (2017) and Reardon et al. (2018), the SEDA approach then estimates a continuous proficiency measure for each state-subject-grade-year and by subgroup using heteroskedastic or homoskedastic ordered probit models. Each state-year-subject has a different mean and standard deviation in order to account for differences in state proficiency examinations over time and across states. These state-level distributions are then placed on the national NAEP performance scale in order to provide an achievement measure that is comparable across over time at the sub-state (district) level.[10] Finally, each subject-grade-year performance distribution is standardized to have a mean of zero and standard deviation of one. Intuitively, these data apply the within-state-year proficiency distributions from state examinations to the cross-state performance measures provided by the NAEP data. Districts in states that perform better on the NAEP examination receive a higher score in the SEDA data, as do districts that perform relatively well on their state's assessment.

[9]Between 2012 and 2017, 44 (45) states changed their math (reading) proficiency metric at least once.

[10]Estimates for cells where the NAEP is not administered (e.g.: even numbered years and grades 3 and 5 through 7) are linearly interpolated and extrapolated.

The data appendix provides a fuller description of this data, and Reardon et al. (2017) and Reardon et al. (2018) provide a more technical treatment.

In total, the SEDA data include approximately 64,000-69,000 district-grade-year math and reading performance observations where at least one school serving grade g participated in CEP at any point through 2017.[11] My main analyses focus on a subset of about 32,000-34,000 district-grade-year observations with baseline district free meal eligibility rates lower than the median among all CEP districts (57.9 percent).

4.3 Other data

The SEDA achievement data is linked to a rich set of baseline area economic and demographic characteristics from the 2006-2010 American Community Survey (ACS). I merge these data to county unemployment rates and per-capita income maintenance payments and district school-aged poverty rates and expenditure composition in order to account for additional time-varying area and school characteristics that might affect student performance and CEP participation.

4.4 Empirical Strategy

In order to examine how CEP affected student academic achievement, I estimate a panel weighted least squares (WLS) differences-in-differences specification, comparing districts that adopted CEP at different points between 2012 and 2017.[12] Districts in which no school chose to participate in CEP are excluded from the analysis. Among ever-participating districts, whether a district is treated in a given year depends on the state in which it is located and the first year any school serving grade g adopted CEP. Districts that first adopted in

[11]The SEDA data reports performance metrics for cells containing at least 20 assessment observations in each group. For example, black achievement measures are only available for district-grades with at least 20 black students; white-black gaps are only available for districts in which there are at least 20 white and 20 black students. Racial gaps are measured according to the standardized mean difference between the distributions for each race/ethnic group.

[12]Following the recommendations in the SEDA documentation, all performance outcomes are weighted by the inverse of the squared standard error of the mean. Columns (5) and (6) of Appendix Table 5 shows the main math results are robust to unweighted and student-enrollment-weighted models.

2016 or 2017 are treated for zero years, while districts in Illinois, Michigan, and Kentucky that adopted the first year of the pilot period are treated for four years. I estimate results separately for math and reading performance with the specification:

$$y_{dgt} = \beta CEP_{dgt} + X'_{dgt}\gamma + \theta_g + \theta_d + \theta_t + \epsilon_{dsgt} \tag{1}$$

Where y_{dsgt} is the achievement score in district d in grade g at time t, expressed in standard deviation units. My preferred specifications focus on a dichotomous treatment where CEP_{dgt} is equal to 1 if any school serving students in grade g in district d participated in CEP in year t.[13] X_{dct} is a vector of time-varying district-grade characteristics that may be correlated with either student performance or district-level decisions to participate in CEP, including the fraction of students who are Hispanic, black, or English-learners in the district; the fraction of Hispanic and black students attending CEP schools; racial and ethnic dissimilarity indices measuring segregation patterns; the student-teacher ratio; county unemployment rates; whether the district is located in a state that is CEP-eligible in year t; and district child poverty rates. Finally, θ_g, θ_d, and θ_t are vectors of grade, district, and year fixed effects, respectively, accounting for factors that do not change within a district or grade over time, and factors that change over time, but affect all states. For example, time fixed effects account for changes in school meal nutritional requirements that applied to all states in 2013. The main analyses stack all grades in order to maximize sample size and statistical power. In sensitivity analyses, I explore whether benefits are concentrated among younger or older students.

Since the sample is limited to districts that participated in CEP by 2017, a causal interpretation of these results requires that the timing of CEP participation is uncorrelated with potential performance, conditional on fixed district factors and time-varying observable characteristics. This assumption would be violated if pilot states were chosen based on potential benefits of CEP adoption, or if districts chose to implement CEP at a point that was

[13] Column (5) of Appendix Tables 5 and 6 shows smaller and less precise results when defining treatment as the share of students in a CEP school. In Section 5.3, I show this pattern is due to high rates of partial participation among districts with low baseline FRP eligibility rates.

most advantageous to student performance. Both policy details and baseline characteristics can inform the plausibility of this identifying assumption. In addition, Section 5.2.2 formally explores this hypothesis with an event study approach.

From a policy perspective, legislation limited the number of pilot states to three in 2012 and four in 2013 and 2014. In determining which states were selected, the Secretary of Agriculture was instructed to "select states with an adequate number and variety of schools and [districts] that could benefit from [CEP]" (Public Law 111-296). In determining the 2012 pilot states, USDA identified states with the greatest number of schools that were likely to qualify based on SNAP participation rates, and allowed ten states to apply (USDA 2011).[14] The selection criteria changed the following two pilot years: all states could apply and states were chosen based on knowledge and awareness of CEP procedures and likely take-up (USDA, 2012, 2013). Baseline academic performance was not a formal criterion in selecting the pilot states, and of the seven states that were eligible to apply but were not selected in 2012, only DC was subsequently chosen as a pilot state.

Examining district baseline characteristics can also suggest whether the timing of CEP participation is correlated with factors that may affect changes in student performance. Figures 4, 5, and 6 display baseline (2009-2011) area economic and district characteristics by year of CEP adoption. In each figure, the solid line shows the distribution of districts that adopted CEP prior to 2016; the dotted line shows the distribution of districts that adopted in 2016 or 2017; and the dashed line shows the distribution of districts that were not participating in CEP as of 2017. These figures show districts with at least one CEP-participating school are more disadvantaged than districts with no participation: prime-age labor force participation rates and median income are lower, and child poverty, income inequality, baseline FRP eligibility, and unemployment rates are higher. Looking at student characteristics, CEP districts tend to have larger shares of black and Hispanic students, and worse academic performance. Differences between early- and late-adopting districts, however, are more muted, suggesting early-adopting districts are more similar to late-adopting areas than

[14] These states were Alaska, DC, Illinois, Kentucky, Louisiana, Michigan, Mississippi, Oklahoma, South Carolina, and Tennessee.

never-adopting districts.

The differences in Figures 4, 5, and 6 suggest CEP districts are not randomly selected and motivate restricting attention to ever-participating districts. Table 2 explores whether area and district characteristics vary among CEP districts across year of implementation.[15] While there are some notable differences – in particular, the initial CEP cohorts have fewer Hispanic students, and the 2012 cohort is more disadvantaged, economic conditions are not differently trending for the earliest adopters and results are robust to excluding districts that adopted the first pilot year (results available upon request).

While these details suggest much of the timing of CEP eligibility was orthogonal to student performance trajectories, it is possible that states with the greatest awareness of the program and relatively well-organized state efforts were selected earlier. If state organization or activity is correlated with both pilot status and achievement trends, leveraging only variation in state-level eligibility timing would lead to biased results. On the other hand, if pilot selection was unrelated to factors shaping student performance, but districts participated in CEP in response to potential student benefits, a participation-based treatment definition will be biased. In practice, disentangling pre-eligibility trends from secular trends is challenging in this setting, as there are only four eligibility "waves" with the vast majority of districts becoming eligible in 2015. With this caveat in mind, Data Appendix Figure 2 suggests that states became eligible for CEP during a period coincident with worsening performance. On the other hand, leveraging both within- and across-state variation allows me to incorporate a rich set of state-specific trends and controls in order to account for state-level factors shaping participation decisions.

To evaluate and account for factors that may affect both the timing of participation and student achievement, I conduct three complementary analyses. First, the main empirical approach controls for all time-invariant district characteristics, as well as many time-varying observable factors that are correlated with CEP participation and performance – such as child poverty rates, the unemployment rate, and the racial/ethnic composition of schools

[15] The 2016 and 2017 adoption years are combined for brevity, as districts adopting in each of these years are untreated throughout the analysis period.

and districts. Second, Section 5 presents event study analyses indicating that there are no significant pre-trends in academic performance for black and Hispanic students prior to CEP adoption after conditioning on district characteristics. Third, I test robustness to a series of standard modifications and extensions, such as exploiting only within-state variation or including linear time trends in baseline variables, following the approach of Hoynes and Schanzenbach (2009) and Hoynes et al. (2016). Findings for math performance are robust to each of these extensions.

5 Results

Only districts with high FRP eligibility are able to participate in CEP. In particular, participating districts must have at least one school with an ISP (and thus, baseline FRP eligibility rate) of at least 40 percent. In practice, many participating districts have baseline eligibility rates well above the minimum threshold: on average, about 58 percent of students were eligible for free meals before CEP, and in 10 percent of CEP districts, more than 80 percent of students were eligible (Figure 7). Recall that the switch to CEP did not substantially change free meal access in districts with relatively high baseline eligibility rates as most students were already eligible for free meals. On the other hand, districts on the eligibility cusp – those with a FRP rate just above the 40 percent ISP threshold – saw free meal access increase up to 60 percentage points under CEP. Therefore, any treatment effect – in terms of both free meal consumption and performance – is likely largest in districts and schools with relatively low baseline eligibility. To examine heterogeneity by the effective treatment "dose", I partition the sample of CEP-adopting districts at the median baseline eligibility rate (57.9 percent). Districts with a baseline eligibility rate less than 57.9 percent form the "exposed" subsample for which CEP led to the largest increases in free meal access.

5.1 CEP and School Meal Participation

In order to establish that CEP affected meal consumption, I collect administrative school-level meal count data in six of the eleven states that adopted CEP before 2015: Georgia, Illinois, Kentucky, New York, Maryland, and West Virginia. Data availability varies by state, and in total, the meal participation data cover approximately 18,800-20,000 school-year observations spanning 2009 through 2016. I merge the meal count data to enrollment information from the Department of Education's Common Core of Data and school-level CEP participation in order to obtain a per-student measure of consumption before and after CEP adoption.

This paper is the first to provide a direct measure of meal consumption for multiple states. One important limitation, however, is that meal count data are not available for all states. I therefore supplement the consumption analyses with information on federal funding districts receive for the school meals program from the Department of Education's School Finance Survey. While the finance data is available for every district in the country, one noteworthy shortcoming is that the reported revenue amount conflates changes in the quantity of meals with changes in the per-meal subsidy rate, both of which are expected to change under CEP.

Table 3 presents the main consumption results from estimating the panel differences-in-differences specification in Equation 1 (at the school level for meal consumption, and district level for federal nutritional assistance funding). Column (1) indicates CEP increased the number of breakfasts served among all CEP schools by 20 meals a student a year (about 38 percent). The change among schools in the exposed district subsample is comparable in both number of meals served and the proportional increase (column (2)). Columns (3) and (4) show the number of lunches increased by 12-13 per student a year for both samples (about 12 percent). Consistent with CEP increasing meal consumption, per-student federal school meal revenue increased by approximately 9 percent (columns 5 and 6). While I find the changes in per-student meal consumption and federal reimbursement are similar in the exposed subsample to the full sample of schools, not all states maintain breakdowns by subsidy rate, and I am unable to fully decipher whether these patterns are due to increases in the number of free meals offsetting reductions in the number of paid meals, or what types of students increase their meal consumption.

In the case where schools and districts adopt CEP in response to increased student demand for school meals, the differences-in-differences regression results in Table 3 would overstate the effect of CEP on meal consumption. In order to investigate whether these findings are the continuation of longer-term trends in school meal participation, as well as how participation evolves after implementation, Figure 8 displays an event study analysis taking the form:

$$y_{spt} = \sum_{p=-5}^{2} [\beta_p \mathbb{1}(P_{spt} = p)] + X'_{spt}\gamma + \theta_s + \theta_t + \varepsilon_{spt} \tag{2}$$

for annual per-student meal consumption y_{spt} in school s p years after the first year of CEP adoption in calendar year t. $\mathbb{1}(P_{spt} = p)$ are a series of indicator variables p years after the first year of implementation; β_p traces out changes in school meal consumption for the full event window, with the year before CEP implementation, β_{-1}, normalized to zero.

Panel (a) shows that for the exposed subsample the number of breakfasts per student was not significantly trending before CEP implementation, and discretely jumped by about 10-20 meals a student a year once CEP was offered.[16] Results for school lunches show that schools tended to implement CEP after lunch participation had been increasing for several years, suggesting that schools may have responded to increasing demand by expanding access to the entire student body. Importantly, however, parametric event studies show a large and strongly significant trend break coinciding with the year of CEP adoption for both breakfasts and lunches, and all samples and specifications (Appendix Table 1). The estimated increase in lunch participation is about 10 meals per student per year (columns 5-8), only slightly smaller than the differences-in-differences results in Table 3.[17]

[16]For the full sample, the pre-period coefficients are jointly significant at the 10 percent level, but both parametric and non-parametric event studies show a discrete increase of at least 10 meals following CEP implementation.

[17]Appendix Table 2 augments the differences-in-differences results with state-specific trends. Under this approach, increases in breakfast consumption are somewhat attenuated (12-13 meals per student), and lunch consumption is similar to the main results in Table 3.

5.2 CEP and Academic Performance

5.2.1 Achievement results

Although greater access to free school meals increased school breakfast and lunch receipt, this consumption may not translate into changes in academic performance as the existing literature finds mixed results of the traditional meals program on academic performance. Schoolwide free meal programs tend to yield more systematic benefits, but these results are somewhat sensitive to how meals are provided and the population studied.

Starting with the full sample of all CEP-participating districts, Table 4 shows CEP did not improve overall reading or math performance. Column (1) estimates Equation 1, including district, cohort, and year fixed effects, but without controlling for time-varying district or economic conditions. Columns (2)-(5) add these characteristics and examine performance among racial/ethnic subgroups. Across specifications and subgroups, there is no significant improvement in math or reading performance.

As the effective treatment "dosage" under CEP depends on a district's baseline eligibility and participation in the free meals program, with the highest-poverty districts experienced little effective change under CEP, the remaining analyses focus on the "exposed" subsample of districts with baseline eligibility rates below the CEP sample median of 57.9 percent. CEP increased free meal access an average of 32 percentage points in these districts, substantially higher than the 23 percentage point increase for lower-exposure (higher baseline eligibility) districts.[18]

The subset of exposed districts differs in several important respects from CEP districts with higher baseline FRP eligibility, summarized in Table 5. First, by definition, these districts had fewer students eligible under the traditional program than non-exposed CEP districts. Similarly, economic conditions – measured by median income, child poverty, and unemployment rates – are slightly better in the exposed subsample. Second, exposed dis-

[18]While this increase in access is larger than the change in the number of lunches served from the meal participation data, recall that the consumption data includes free, reduced, and paid meals. Under CEP, some students who previously purchased a school meal would continue to receive a school meal, but would have no out-of-pocket costs.

tricts have slightly smaller minority shares, and had higher math and reading performance prior to CEP. Third, in exposed areas, participation decisions are less likely to be made at the district level. Whereas 70 percent of non-exposed participating districts had full district participation, 57 percent of exposed districts fully participated by 2017. All of these patterns suggest the exposed sample consists of relatively low-poverty districts (compared to other CEP districts. In an absolute sense, even low-poverty CEP districts tend to be more disadvantaged than non-participating districts).

Focusing on the subsample of exposed districts with the largest gains in access to free meals indicates important treatment heterogeneity: for this group, CEP improved overall math performance by about 0.02 standard deviations (Table 6, column (1)). Scaling the this intent-to-treat estimate by the fraction of students gaining access to free meals (32 percent) implies that free school meals improve overall math performance by approximately 0.05 standard deviations.

These improvements in math performance are concentrated among populations that are especially likely to gain access to free meals under CEP. In particular, income-eligible Hispanics have lower participation rates than other groups (Chaparro et al., 2014; Goerge et al., 2009), while white students tend be less likely to qualify on the basis of income. The remaining columns of Table 6 show Hispanic math performance by approximately 0.03 standard deviations (column (3)), and white math performance about 0.02 standard deviations (column (4)). In contrast, math performance among black students did not significantly change. Therefore, column (5) suggests while white-black performance gaps widened following CEP, this measure is driven by improvements in the absolute performance of white students, rather than deteriorating performance among black students.

Panel (b) presents analogous results for reading. Across subgroups, reading outcomes are smaller in magnitude and less precisely estimated. A series of robustness checks also demonstrates the magnitude and sign of reading performance is sensitive to the specification and sample. Given these patterns, the remaining analyses focus on math performance; reading results are included in the Appendix for completeness.

The reduced-form estimates in Table 6 provide the intent-to-treat (ITT) effect of offering free meals to all students. Recovering the effect of actual meal consumption on treated students (TOT) is not feasible with district (or school) aggregate performance measures. In principle, instrumenting the change in meal consumption by CEP participation would obtain the district-level TOT. In practice, however, meal data are only available for schools in six states, and only half of these observations contain breakdowns by payment category. To overcome these data limitations, I consider how the change in federal funding for school meals induced by CEP participation affects student performance.

Appendix Table 3 reports results from instrumenting per-student federal school meal spending with CEP participation, and shows that an additional $1,000 per student in school meals induced by CEP adoption increases math scores by an insignificant 0.16 standard deviations for the full sample (column 1) and 0.51 standard deviations for the exposed subsample. Reading does not significantly improve in either sample. As the average district received approximately $100 in additional per-pupil school meal reimbursement under CEP, this implies access to schoolwide free meals improved math scores by 0.05 standard deviations, consistent with the scaled results in Table 6. Interestingly, while average changes in federal funding are similar for both samples (about $100), improvements are concentrated in the exposed subsample. This pattern suggests that not only revenue amounts, but also which types of students benefit from additional resources, is important for understanding changes in average performance.

5.2.2 Timing of CEP adoption

The differences-in-differences specifications provide an estimate of the average effect of CEP one to four years after implementation, relative to previous years. One outstanding question is how districts were performing prior to CEP, and whether any benefits grew or diminished with program experience. To this point, event study analyses can illustrate the extent to which student performance changed over the full analysis period, and whether CEP adoption coincides with previous trends in student performance. Figure 9 displays event study plots for math performance by estimating Equation 2 at the district level. These plots show math performance was not significantly trending for black (panel b) or Hispanic (panel c) students prior to CEP adoption. However, CEP adoption followed a period of improvements in white students' math performance (panel d). Reflecting the fact that white students account for a large share of the total student population, there is suggestive evidence of improvements in overall performance over the longer term (panel a). Appendix Figure 1 shows these general patterns are similar after including state linear trends and trends in baseline variables. In contrast, reading performance does not display any significant pre-treatment (or post-treatment) pattern for any subgroup (Appendix Figure 2). If anything, black performance was slightly worsening prior to CEP adoption.[19] Given these patterns, I emphasize results for white students should be interpreted cautiously, as districts tended to adopt CEP during a period of secular improvements in performance. In contrast, the timing of CEP implementation is more plausibly exogenous from the perspective of black and Hispanic math performance.[20]

The analyses in Figures 9 and Appendix Figure 2 display unbalanced event studies, binning all years earlier than -5 (years -8 through -5 for the 2017 cohort) and years later than 2 (years 2 and 3 for the 2012 cohort) in order to provide suggestive evidence the extent to which benefits from schoolwide free meals cumulate or diminish over time. Although point estimates generally suggest greater achievement gains with each year of access, I cannot reject equal treatment effects in each of the first three years of participation.[21] As more districts gain experience with schoolwide meals, greater exploration of this topic can inform the extent to which there are diminishing marginal returns to each year of access.

[19]Patterns are similar with the inclusion of state and baseline variable trends.

[20]Data Appendix Table 2 shows that limiting the source of identifying variation to the year of eligibility is likely problematic in this setting. Both math and reading performance were steadily and significantly worsening for most subgroups before their state became eligible to adopt CEP. These patterns call into question the exogeneous nature of eligibility decisions, at least with respect to student performance, but due to the short nature of the phase-in period, it is difficult to disentangle these patterns from secular trends in performance.

[21]While only the earliest and latest-adopting districts contribute to the tails in these figures (and these districts account for a small share of all CEP districts), balanced event study plots show qualitatively similar patterns (Appendix Figure 3).

5.2.3 Extensions: Heterogeneity and Sensitivity analyses

Performance by grade level Previous work has found that universal meal programs have particularly large benefits for young children (Gordanier et al., 2019; Fuller and Comperatore, 2018; Gordon and Ruffini, nd). Consistent with the existing literature, Table 7 suggests that CEP improved math performance significantly more for black and Hispanic elementary (grades 3-5) than middle (grades 6-8) schoolers.

District resources and student composition By changing the federal school meal reimbursement formula, CEP adoption may have affected performance by altering district revenue or resources, and any changes in district resources may contribute to changes in academic performance. Table 8 explores this possibility by examining various resource measures and shows that total federal revenues (column (1)), federal revenues net of nutritional assistance payments (column (2)), and total per-pupil expenditures (column (3)) did not significantly change following CEP adoption. Columns (4) and (5) show that per-pupil instructional expenditures and student-teacher ratios slightly fell, suggesting districts increased the number of instructional staff at lower salaries.[22] These changes in resources, however, do not drive the main findings. When controlling for per-student total and instructional expenditures, results are nearly identical to the baseline results in Table 6 (Appendix Table 5, column (1)).

Since free school meals provide an in-kind subsidy to families (with a fungible value of about $4.50 a school day), CEP adoption may have changed the student composition of a district if families moved into adopting districts or transferred from private schools to public schools. To explore whether CEP changed district composition, Table 9 panel (a) regresses district-level student characteristics on CEP participation. These results show no changes in the share of black or white students enrolled in a district following CEP implementation, and a slight reduction in the share of Hispanic students (0.3 percentage points, or about 2 percent).

[22] In additional results, I find no statistically or economically significant change in district-grade enrollment.

Even with minor shifts at the district level, CEP could prompt intra-district sorting or changes in segregation patterns if districts realigned school boundaries in order to maximize CEP eligibility or if students transferred in or out of CEP schools based on perceived benefits. Such sorting is mostly likely to occur in districts where some, but not all, schools adopt CEP. Panel (b) of Table 9 examines school-level demographic shifts for the subset of districts in which at least one school in the district participated in CEP and at least one did not by measuring the fraction of students in each racial/ethnic group attending CEP schools. In these districts, there is no economically or statistically significant change in the fraction of students from any racial/ethnic group attending a CEP school. Looking more directly at segregation patterns, panel (c) suggests CEP slightly increased the concentration of white and black students in a school, measured by district-grade dissimilarity indices.

All of the previous analyses controlled for segregation patterns and student demographics at both the district and school level in order to account for shifts in student composition. As a complementary exercise to place bounds on the extent that changes in the student population can account for observed changes in performance, Appendix Table 4 defines predicted performance for each subgroup-subject as the grade-specific fitted values from regressing district segregation and student composition on performance. While point estimates suggests demographic shifts coinciding with CEP would lead to improved performance for Hispanic and white students, these demographic changes can explain no more than ten percent of the observed improvements in math performance.

Alternative specifications and sample definitions Appendix Table 5 explores the robustness of improvements in math performance. As mentioned previously, results are unchanged when including controls for financial resources, suggesting that changes in district resources are not driving the main results (column (1)). Columns (2) and (3) add state-specific linear trends and trends in baseline child poverty, unemployment, student racial/ethnic composition and segregation, and student-teacher ratios in order to account for possible performance trends that are correlated with the timing of implementation, following the approach in Hoynes et al. (2016). Column (3) additionally controlls for prior-year performance. While effects for black and white students are sharply attenuated under this approach, estimates for Hispanic students are relatively unchanged. Column (4) includes state-by-year fixed effects, thereby only exploiting variation in CEP adoption within a state. This specification effectively pools many (51) single-state analyses, and again, results point to improvements in math performance, particularly among Hispanic students.

During the 2012 through 2017 period, about 60 percent of participating districts had full participation, and more than 75 percent of students attended a CEP school in an additional 13 percent of districts. Given this distribution, the main results define treatment as a binary indicator, regardless of the share of students attending CEP schools. An alternative approach would define a district's "treatment" as the share of students attending CEP schools. Under this approach in Appendix Table 5 column (5), results are smaller in magnitude and much less precise, but confidence intervals cannot rule out improvements of the magnitude shown in Table 6. The differences between estimates using a binary and continuous treatment variable suggest districts with partial participation experienced the largest benefits. I return to this issue in Section 5.3.

Finally, columns (6) and (7) return to the main estimating equation (Equation 1), but instead of implementing weighted least squares, these columns weigh the results either equally across districts (column (6)) or by the log of baseline student enrollment (column (7)). In both cases, estimates are similar to the main findings, suggesting that any benefits are not concentrated in particularly large or small districts. Appendix Table 6 presents the corresponding results for reading performance. Here, findings are less consistent across specifications. When even a parsimonious set of additional controls are included, I cannot rule out no effect of CEP on reading performance for any subgroup.

I also explore robustness to different sample definitions in Appendix Table 7 (math) and 8 (reading). As each of these alternative samples is small relative to the main results, the loss of statistical power precludes making definitive conclusions and these results should be viewed as suggestive. Column (1) limits the sample to districts with full participation – districts in which the binary treatment measure coincides with the fraction of students attending a universal free meal school. Moreover, student sorting across schools in response to CEP is less likely to occur in these districts. Although point estimates are attenuated, confidence intervals cannot rule out changes in performance comparable to those reported in Table 6. Column (2) focuses on the balanced panel of districts that have a valid observation for each year covered in the SEDA data to assuage concerns that the main findings – particularly for race/ethnic groups – are driven by changes in the number of students of a race/ethnic group enrolled in the district, or by the inclusion of states that experienced changes in their state examinations over this period.[23] Again, results point in the same direction as the main results. Column (3) lastly limits the sample to districts that adopted CEP the first year their state became eligible. In these districts, CEP timing is driven by statutory eligibility, rather than district-level decisions. While these estimates cannot reject the null hypothesis that CEP had no effect, they also cannot rule out changes of the magnitude reported in Table 6. Appendix Table 8 again illustrates the sensitivity of the reading results, and in general, there is no systematic evidence CEP improved reading among any student or district subgroup.

Finally, Appendix Figures 4 and 5 plot coefficients and confidence intervals from the specifications in Table 6, but dropping a single Census Division, state, or grade in order to explore whether treatment effects vary by geography or grade level. These figures illustrate that results are not driven by the experiences of a single geographic area, and consistent with Table 7, younger students tend to experience larger math improvements.

5.3 Effects throughout the exposure distribution

The main analyses focus on the set of CEP districts below the baseline free meal eligibility median (57.9 percent). Across the full sample of CEP-participating districts, however, there is wide variation in baseline FRP eligibility: About 12 percent of CEP districts have a baseline eligibility below 40 percent, and 10 percent have baseline eligibility above 80 percent (Figure 7).[24]

[23] Reardon et al. (2018) details the methodology for censoring and excluded observations. The most common reason for exclusion is a substantial change in a state's examination.

[24] The cut variable is defined on district baseline eligibility and does not factor in the share of schools within the district actually adopting CEP in order to allay concerns that groups of schools within districts

Recall that "exposure" is defined as $1 - pctFRP_{dg,2009-2011}$: the share of students in district d, grade g *ineligible* for free and reduced meals between 2009 and 2011. If all schools in a district participate in CEP, exposure is equivalent to the share of students gaining access under CEP. This is the case for about the 60 percent of CEP districts with full participation. In districts with partial participation, the share of students gaining access under CEP is less than a district's exposure. In addition, the likelihood that all schools within a district participate is increasing in district baseline FRP eligibility, illustrated in Figure 10, panel (a). Accordingly, while the share of students with access to free meals (either through the traditional formula or CEP) is increasing in baseline eligibility (panel (b)), there is an inverse-U shape relationship between baseline eligibility and the fraction of students *gaining* access under CEP – the effective "treatment" dosage. Figure 10 panel (c) shows that access gains are largest for districts with baseline poverty shares between about 50 and 60 percent. The highest-exposure schools (those with the lowest baseline FRP eligibility) experience relatively little effective treatment under CEP as relatively few schools in the district participate; and the lowest-exposure schools also experience essentially no gains due to high baseline eligibility.

Mirroring the distribution of increased eligibility, there is a non-monotonic relationship between exposure to CEP and performance gains, both in the the aggregate and for Hispanic students, shown in Table 10. Although not statistically different across the distribution, point estimates suggest CEP conferred the largest benefits to Hispanic students in districts with less than 50-60 percent of students eligible for free meals at baseline. In contrast, any improvements in black math performance are concentrated in districts with the lowest baseline eligibility rates.

Appendix Table 9 presents analogous results for reading. While point estimates suggest that any improvements are limited to the highest-income districts (those with baseline eligibility rates below 40 percent), none of these findings is statistically different from zero.

One possible explanation for these patterns is that the *students* gaining access in rela-

may strategically apply for CEP in order to maximize total revenue or participation.

tively high-exposure districts had the largest marginal benefits from additional nutritional assistance. While this hypothesis is untestable with district-level data, evidence from state studies suggests there may be heterogeneity across areas in the types of students benefiting the most from universal provision, although the existing work has not reached a consensus on this point. In the case of South Carolina, low-income students who did not receive TANF or SNAP experience the largest improvements (Gordanier et al., 2019), while New York City students who were previously ineligible for free meals benefited the most from universal provision (Schwartz and Rothbart, nd).

5.4 Heterogeneous effects

Besides the share of students previously eligible for free meals, the benefits of schoolwide free meals may vary with other area characteristics. Unlike programs like SNAP or TANF that provide a near-cash benefit, school meals are a quantity-based form of assistance. While the monetary value of in-kind benefits is higher in expensive areas, the additional resources from free meals increase family purchasing power by a greater amount in low-cost of living areas. With decreasing marginal benefits of consumption, additional nutritional assistance is also expected to be higher in areas where few families receive other income assistance programs (conditional on income). Finally, with the caveat that there are few urban areas in each state, several single-state papers have documented CEP tends to yield greater benefits in non-urban areas (Gordanier et al., 2019; Fuller and Comperatore, 2018).

To explore whether the benefits of schoolwide meals are concentrated in any of these areas, I partition the main analysis sample based on urban location and at the CEP-sample baseline median cost-of-living, SNAP receipt, and per-capita income assistance levels. Table 11 suggests that math performance improvements are concentrated in areas with relatively low costs of living, consistent with CEP providing greater purchasing power for families in these areas. More granular student-level data that includes measures of family income and consumption could provide more concrete evidence on this hypothesis. In contrast, there are no systematic differences in treatment effects by urban location or participation in other income assistance programs (columns (2) through (4)).

6 Discussion and conclusion

This paper finds that schoolwide free meal programs increase breakfast and lunch participation. In addition, schoolwide free meals led to modest improvements in math performance among groups likely to gain access to free meals under universal provision, as well as for younger students. Results are not driven by concurrent changes in school resources or observable features of the school environment.

These findings are largely consistent with results from papers examining the effect of CEP in a single state, as well as findings from earlier universal breakfast programs. In South Carolina, Gordanier et al. (2019) find that CEP improved math performance among elementary by about 0.06 standard deviations, with smaller effects in reading and among middle school students. In North Carolina, Fuller and Comperatore (2018) find elementary math performance improved approximately 0.02 to 0.03 standard deviations, middle school math performance did not significantly improve, and both reading performance improved approximately 0.04 standard deviations for both grade levels. Schwartz and Rothbart (nd) find improvements of a similar magnitude (0.03 standard deviations) among students who qualified for free meals under the traditional program, and larger improvements among higher-income, previously-ineligible students. While schoolwide free meals did not significantly improve academic performance in other settings (Kho, 2018; Leos-Urbel et al., 2013), my analyses suggest that negligible aggregate effects of such programs may mask differential effects across student populations or schools, depending on the magnitude of the effective "treatment."

Importantly, prior single-state evaluations of CEP compare the experiences of schools that choose to implement CEP with those that do not. In contrast, all districts in my sample opted to participate in CEP, but differ in the timing of adoption. That these studies leverage different sources of variation, yet reach broadly similar conclusions supports an earlier body of work pointing to the role of nutritional assistance and school investments in improving short-term outcomes for students.

When interpreting these results, it is important to recall that CEP expands free meals to two types of students. One group is students who live in a high-poverty district, but whose family incomes are greater than the cutoff for free meals. The second is students who are income-eligible for the traditional program, but who are not receiving other forms of assistance and whose families did not complete the required paperwork. The literature has not reached a consensus of which *students* benefit the most from universal access. Given the aggregate nature of district-level data, I am unable to fully explore the extent to which *individual* benefits differ by receipt of other forms of assistance or family income. Scores by race and ethnicity can provide insights into this heterogeneity only if race/ethnicity is correlated with students' socioeconomic status.

While the modest improvements in math performance documented in this paper are similar in magnitude to other papers examining the effect of CEP and related schoolwide meals programs in a single state, they are small relative to earlier district-led initiatives that modified how school meals were served. In particular, existing work suggests moving breakfasts from before school to during the school day leads to slightly larger improvements than schoolwide free meals (Dotter, 2013; Imberman and Kugler, 2014). However, considering the size and generosity of the implied income transfer, CEP offers benefits similar in magnitude to other forms of income assistance. For example, an additional \$1,000 in EITC payments increases test scores for 3-8th graders by about 0.04 standard deviations (Dahl and Lochner, 2017). By these metrics, CEP delivered benefits on the order of a \$500 family income transfer, for a cost to the federal government of approximately \$100 a student. Taken as a whole, these results suggest that school-based assistance can yield important benefits, particularly for groups unlikely to have access to traditional, family-income based programs.

References

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

Anderson, M. L., Gallagher, J., and Ritchie, E. R. (2017). School lunch equality and academic performance. Technical report, National Bureau of Economic Research.

Arteaga, I. and Heflin, C. (2014). Participation in the national school lunch program and food security: An analysis of transitions into kindergarten. *Children and Youth Services Review*, 47.

Bartlett, S., Klerman, J., Wilde, P., Olsho, L., Logan, C., Blocklin, M., Beauregard, M., and Enver, A. (2014). Evaluation of the Healthy Incentives Pilot (HIP): Final report. *Food and Nutrition Service, USDA. Cambridge, MA: Abt Associates.*

Bhattacharya, J., Currie, J., and Haider, S. J. (2006). Breakfast of champions? the School Breakfast Program and the nutrition of children and families. *Journal of Human Resources*, 41(3):445–66.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415.

Chaparro, M. P., Harrison, G. G., and Pebley, A. R. (2014). Individual and neighborhood predictors of participation in the Supplemental Nutrition Assistance Program (SNAP) in Los Angeles county. *Journal of hunger & environmental nutrition*, 9(4):498–511.

Coleman-Jensen, A., Rabbitt, M. P., Gregory, C., and Singh, A. (2017). Household food security in the United States in 2017. *Economic Research Service, Economic Research Report 256.*

Dahl, G. B. and Lochner, L. (2017). The impact of family income on child achievement: Evidence from the Earned Income Tax Credit: Reply. *The American Economic Review*, 107(2):629–631.

Davis, W. and Musaddiq, T. (2018). Estimating the effects of subsidized school meals on child health: Evidence from the Community Eligibility Provision in Georgia schools.

Domina, T., Pharris-Ciurej, N., Penner, A. M., Penner, E. K., Brummet, Q., Porter, S. R., and Sanabria, T. (2018). Is free and reduced-price lunch a valid measure of educational disadvantage? *Educational Researcher*, 47(9):539–555.

Dotter, D. D. (2013). Breakfast at the desk: The impact of universal breakfast programs on academic performance. *Mathematica Policy Research*.

Dunifon, R. and Kowaleski-Jones, L. (2003). Associations between participation in the national school lunch program, food insecurity, and child well-being. *Social Service Review*, 77:72–92.

Education Week (2014). Big year looms for Common Core testing. `https://www.edweek.org/ew/articles/2014/09/03/03assessment.h34.html`.

Fletcher, J. M. and Frisvold, D. E. (2017). The relationship between the School Breakfast Program and food insecurity. *Journal of Consumer Affairs*, 51(3):481–500.

Food Research and Action Center (2019). Community Eligibility: The key to hunger-free schools, school year 2018-2019.

Frisvold, D. E. (2015). Nutrition and cognitive achievement: An evaluation of the School Breakfast Program. *Journal of Public Economics*, 124:91–104.

Fuller, S. and Comperatore, A. (2018). A free lunch: School meal uptake and student achievement under the Community Eligibility Provision. Association for Public Policy Analysis and Management, Fall Conference.

Ganong, P. and Liebman, J. B. (2013). The decline, rebound, and further rise in snap enrollment: Disentangling business cycle fluctuations and policy changes. *NBER Working Paper 19363*.

Gassman-Pines, A. and Bellows, L. (2018). Food instability and academic achievement: a quasi-experiment using SNAP benefit timing. *American Educational Research Journal*, 55(5):897–927.

Gennetian, L. A., Seshadri, R., Hess, N. D., Winn, A. N., and Goerge, R. M. (2016). Supplemental Nutrition Assistance Program (SNAP) benefit cycles and student disciplinary infractions. *Social Service Review*, 90(3):403–433.

Glantz, F. B., Berg, R., Porcari, D., Sackoff, E., and Pazer, S. (1994). School lunch eligible non-participants. Technical report.

Gleason, P. and Suitor, C. (2001). Children's diets in the mid-1990s: Dietary intake and its relationship with school meal participation. Technical report.

Goerge, R. M., Franzetta, K., and Dilts, J. (2009). Food stamp program eligibility and participation in Chicago. *Chicago, IL: Chapin Hall Center for Children.*

Gordanier, J., Ozturk, O., Williams, B., and Zhan, C. (2019). Free lunch for all! the effect of the Community Eligibility Provision on academic outcomes. Technical report.

Gordon, N. E. and Ruffini, K. J. (n.d). School nutrition and student discipline: Effects of schoolwide free meals. *Education Finance and Policy*, page forthcoming.

Gundersen, C., Kreider, B., and Pepper, J. (2012). The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics*, 166:79–91.

Hinrichs, P. (2010). The effects of the National School Lunch Program on education and health. *Journal of Policy Analysis and Management*, 29:479–505.

Howard, L. L. (2011). Does food insecurity at home affect non-cognitive performance at school? A longitudinal analysis of elementary student classroom behavior. *Economics of Education Review*, 30:157–176.

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–934.

Hoynes, H. W., McGranahan, L., and Schanzenbach, D. W. (2015). Snap and food consumption. *SNAP Matters: How Food Stamps Affect Health and Well Being, eds. J. Bartfeld, C. Gundersen, T. Smeeding, and J. Ziliak, Stanford University Press.*

Hoynes, H. W. and Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–39.

Imberman, S. A. and Kugler, A. (2014). The effect of providing breakfast on achievement and attendance: Evidence from an in-class breakfast program. *Journal of Policy Analysis and Management*, 33.

Institute of Education Sciences (2009). Appendix A: State testing programs under NCLB. `https://ies.ed.gov/ncee/pubs/2009013/appendix_a.asp`.

Kho, A. (2018). *Three Essays on School Reform*. Vanderbilt University.

Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, 116:777–803.

Leos-Urbel, J., Schwartz, A. E., Weinstein, M., and Corcoran, S. (2013). Not just for poor kids: The impact of universal free school breakfast on meal participation and student outcomes. *Economics of Education Review*, 36:88–107.

Logan, C. W., Connor, P., Harvill, E. L., Harkness, J., Nisar, H., Checkoway, A., Peck, L. R., Shivji, A., Bein, E., Levin, M., and Enver, A. (2014). Community Eligibility Provision evaluation. *Prepared by Abt Associates for the USDA, Food and Nutrition Service.*

Mabli, J. and Worthington, J. (2014). Supplemental Nutrition Assistance Program participation and child food security. *Pediatrics*, 133(4):1–10.

Meyers, A. F., Sampson, A. E., Weitzman, M., Rogers, B. L., and Kayne, H. (1989). School Breakfast Program and school performance. *American journal of diseases of children*, 143(10):1234–1239.

Neuberger, Z., Segal, B., Nchako, C., and Masterson, K. (2015). Take up of Community Eligibility this school year. Technical report, Center on Budget and Policy Priorities.

Nord, M. and Romig, K. (2006). Hunger in the summer: seasonal food insecurity and the National School Lunch and Summer Food Service programs. *Journal of Children and Poverty*, 12(2):141–158.

Ratcliffe, C., McKernan, S.-M., and Zhang, S. (2011). How much does the supplemental nutrition assistance program reduce food insecurity? *American Journal of Agricultural Economics*, 97:1082–1098.

Reardon, S. F., Ho, A. D., and Kalogrides, D. (2019). Validation methods for aggregate-level test scale linking: A case study mapping school district test score distributions to a common scale. *Stanford Center for Education Policy Analysis, Working Paper 16-09*.

Reardon, S. F., Shear, B. R., Castellano, K. E., and Ho, A. D. (2017). Using heteroskedastic ordered probit models to recover moments of continuous test score distributions from coarsened data. *Journal of Educational and Behavioral Statistics*, 42(1):3–45.

Reardon, S. M., Rahle, E. M., Shear, B. R., Kalogrides, D., DiSalvo, R., and Ho, A. D. (2018). Stanford education data archive: Technical documentation, version 2.1. `https://stacks.stanford.edu/file/druid:db586ns4974/SEDA_documentation_v21.pdf`.

Schanzenbach, D. W. (2009). Do school lunches contribute to childhood obesity? *Journal of Human Resources*, 44(3):684–709.

Schanzenbach, D. W. and Zaki, M. (2014). Expanding the School Breakfast Program: Impacts on children's consumption, nutrition and health. *NBER Working Paper 20308*.

Schwartz, A. and Rothbart, M. (n.d). Let them eat lunch: The impact of universal free meals on student performance. *Journal of Policy Analysis and Management*, page forthcoming.

US Department of Education (2013). States granted waivers from No Child Left Behind allowed to reapply for renewal for 2014 and 2015 school years. https://www.ed.gov/news/press-releases/states-granted-waivers-no-child-left-behind-allowed-reapply-renewal-2014-and-2015-sch

US Department of Education (2017). President's fiscal year 2018 budget request for the U.S. Department of Education. https://www2.ed.gov/about/overview/budget/budget18/justifications/index.html.

USDA (2002). Provision 2 guidance. http://www.fns.usda.gov/sites/default/files/Prov2Guidance.pdf.

USDA (2012). Community Eligibility Option: Guidance and process for selection of states for school year 2012-2013 memorandum sp 12 -2012. http://www.fns.usda.gov/sites/default/files/SP12-2012%20os.pdf.

USDA (2013). Community eligibility option: Guidance and process for selection of states for school year 2013-2014 memorandum sp 15 -2013. http://www.fns.usda.gov/sites/default/files/SP15-2013os.pdf.

USDA (2017). Congressional justification, fiscal year 2018. https://www.obpa.usda.gov/32fnsexnotes2018.pdf.

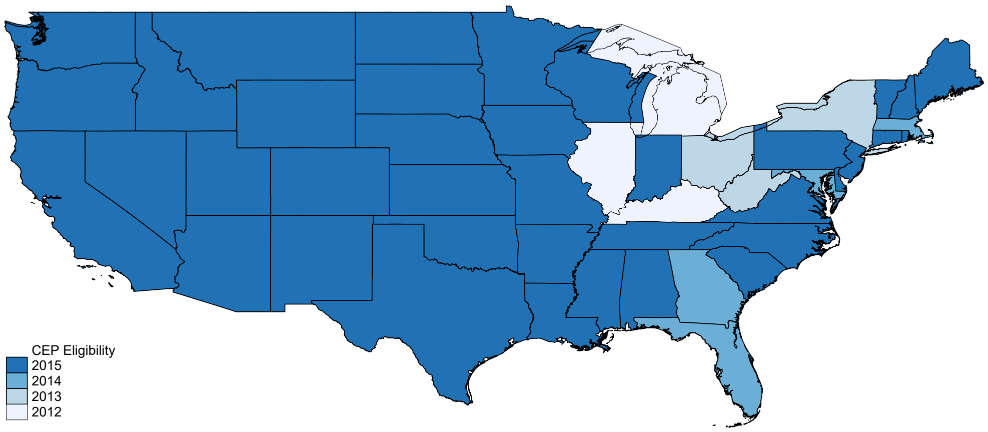
USDA (2018a). Characteristics of SNAP households: Fiscal year 2016. http://www.fns.usda.gov/sites/default/files/ops/Characteristics2014.pdf.

USDA (2018b). National level annual summary tables: Fiscal years 1969-2017. http://www.fns.usda.gov/pd/child-nutrition-tables.

Figure 1: Fraction 5-17 Year-Olds Receiving School Meals by Payment Category

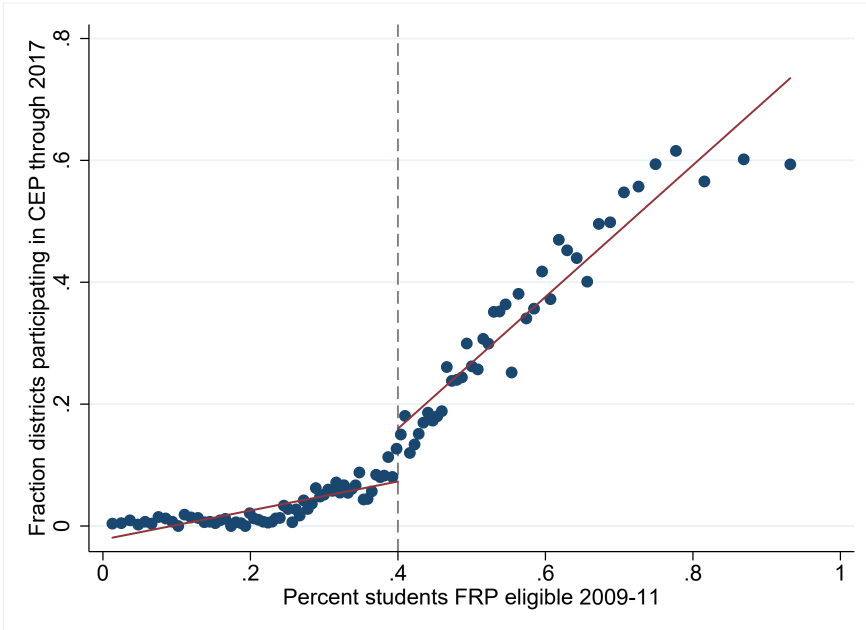
Notes: Figure shows the share of 5-17 year-olds receiving a school meal each year by payment amount. School meal counts from USDA (2018). Population estimates from Census Bureau decennial census and intercensal estimates. Left dashed vertical line denotes 2012, the year schools in the first pilot states became eligible to adopt CEP. Right dashed line denotes 2015, the first year schools in all states were eligible to adopt CEP.

Figure 2: First Year of CEP Availability, by State



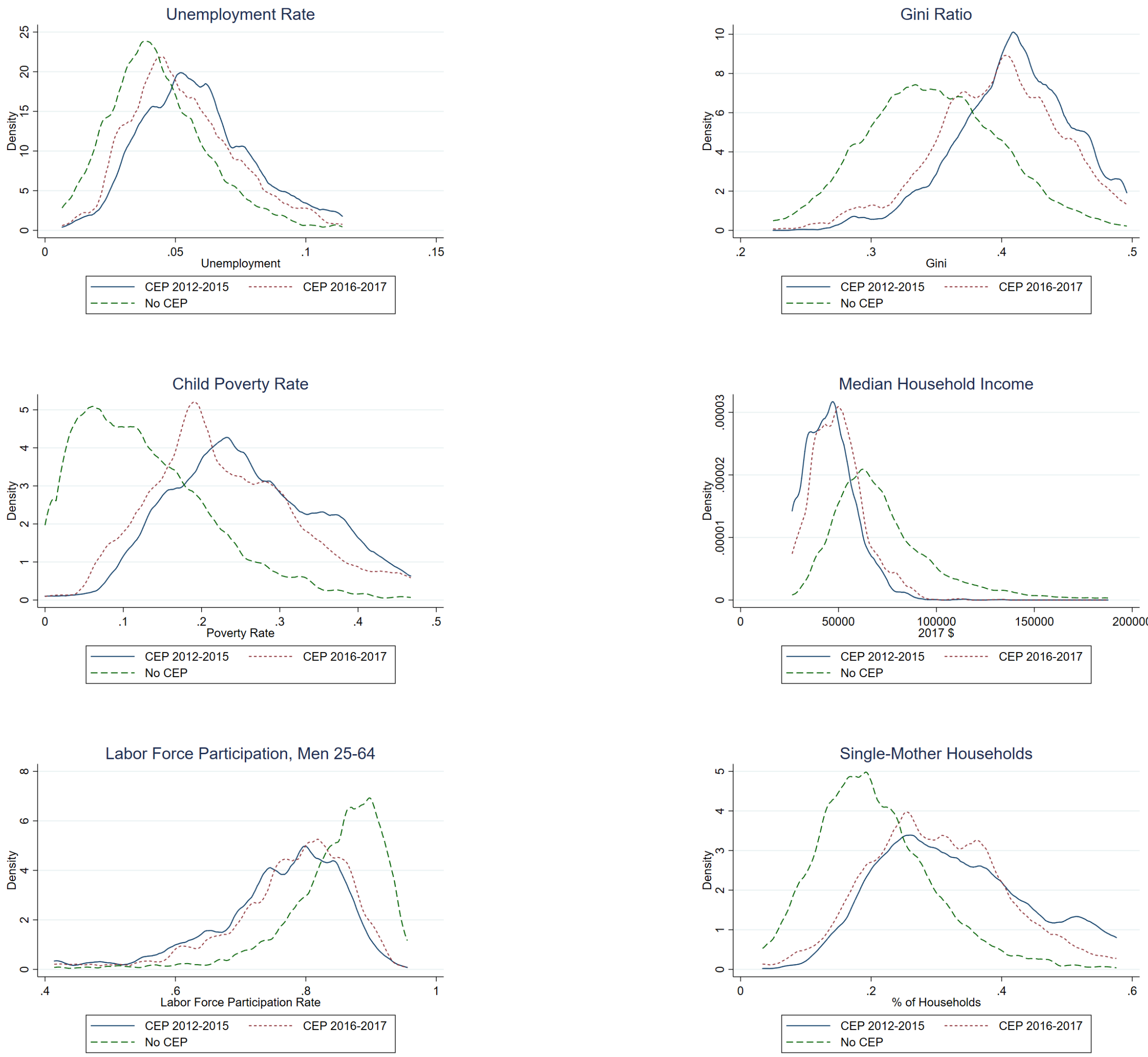
Notes: Source USDA, 2011, 2012, 2013, 2014.

Figure 3: CEP Participation by Baseline Share of Students Eligible for Free Meals



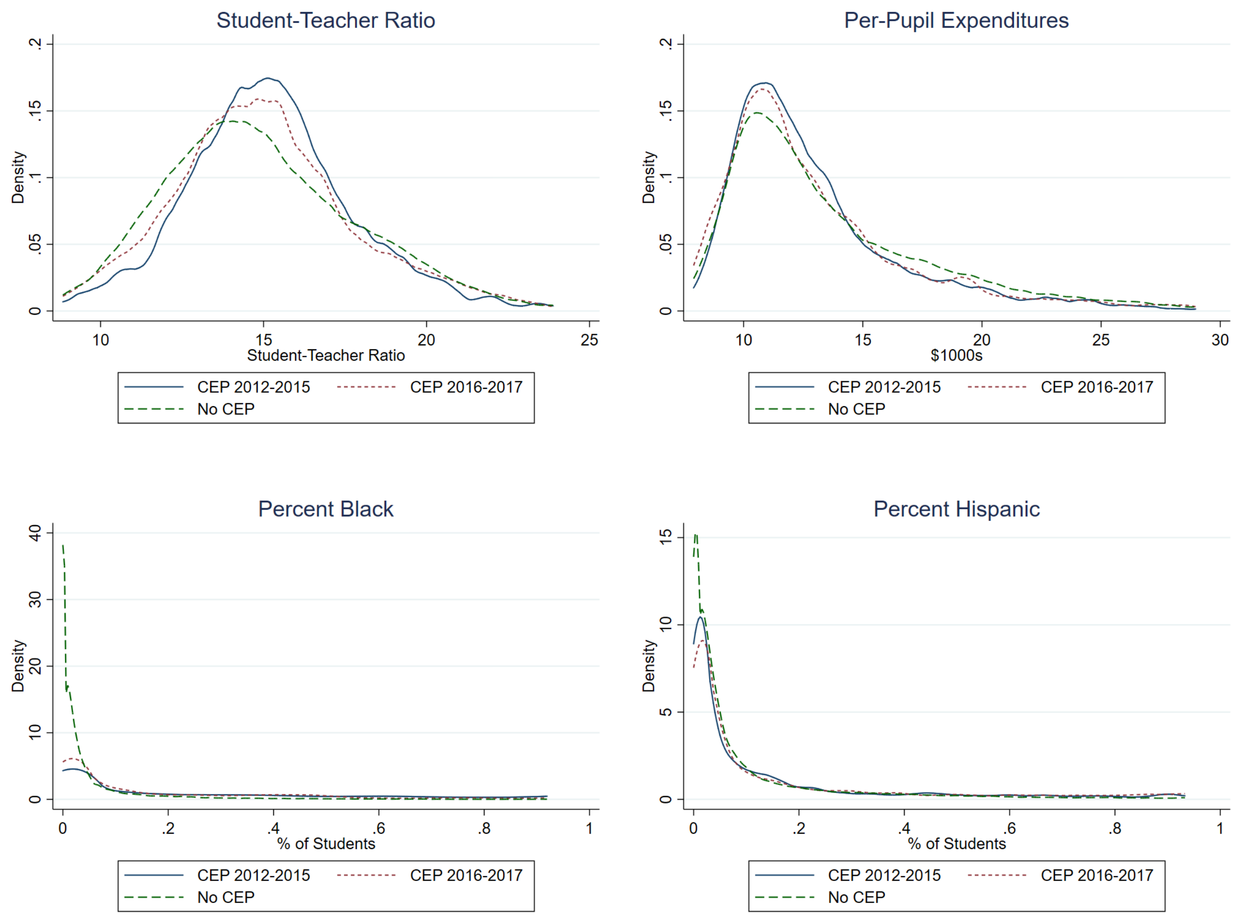
Notes: Figure plots relationship between baseline share of FRP students in a district in 2009-2011 (horizontal axis) and the probability a district participated in CEP by 2015 (vertical axis) from a binscatter of 100 equal-sized bins. The vertical line at 40 percent shows the minimum (school-level) FRP eligibility rate required for CEP participation. Sources: USDA FRAC/CBPP, Common Core of Data.

Figure 4: Baseline Area Economic-Well-being by CEP Adoption



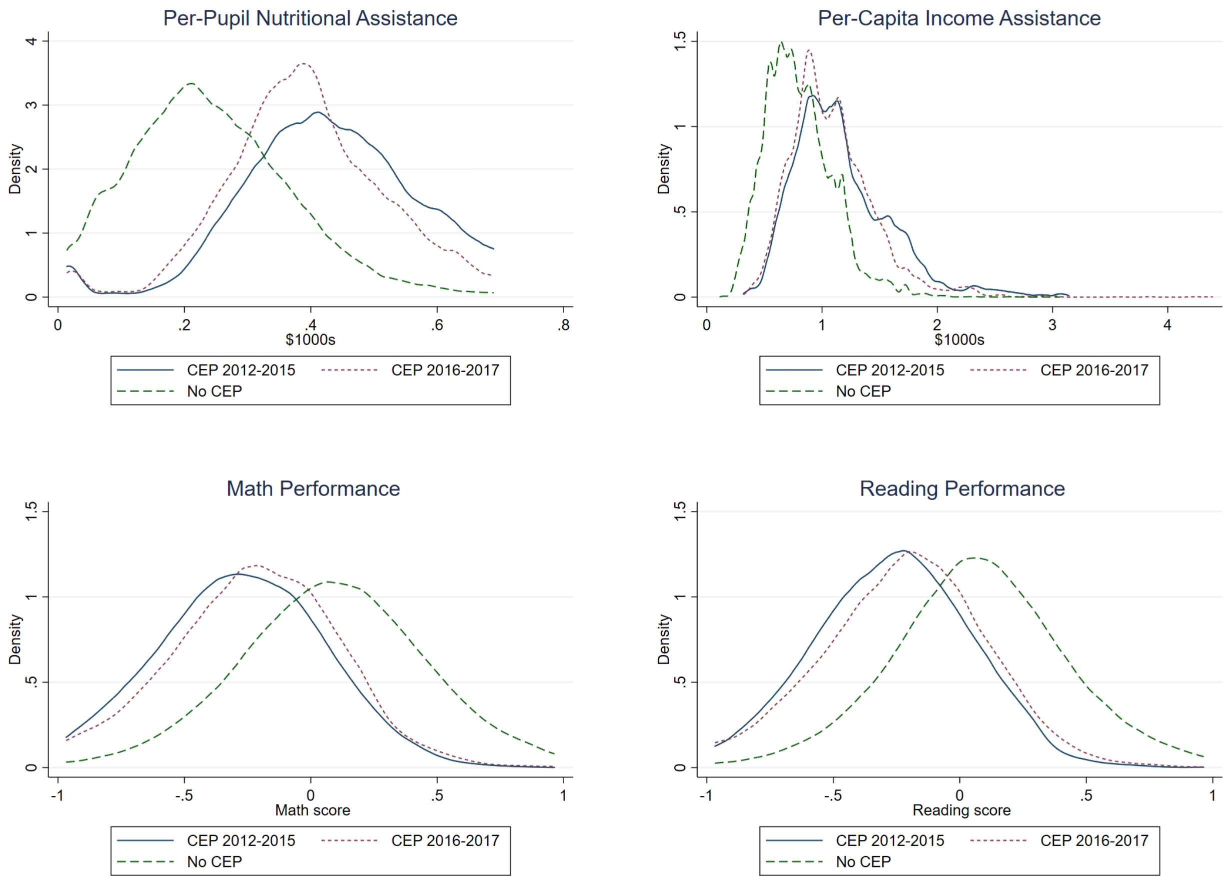
Notes: Figure shows the distribution of baseline area characteristics (2006-2010 for ACS variables, 2009-2011 for unemployment and labor force participation) by the year a district-grade first participated in CEP. "No CEP" are district-grades that did not adopt CEP by 2017. See text and data appendix for variable details.

Figure 5: Baseline District Student Characteristics by CEP Adoption



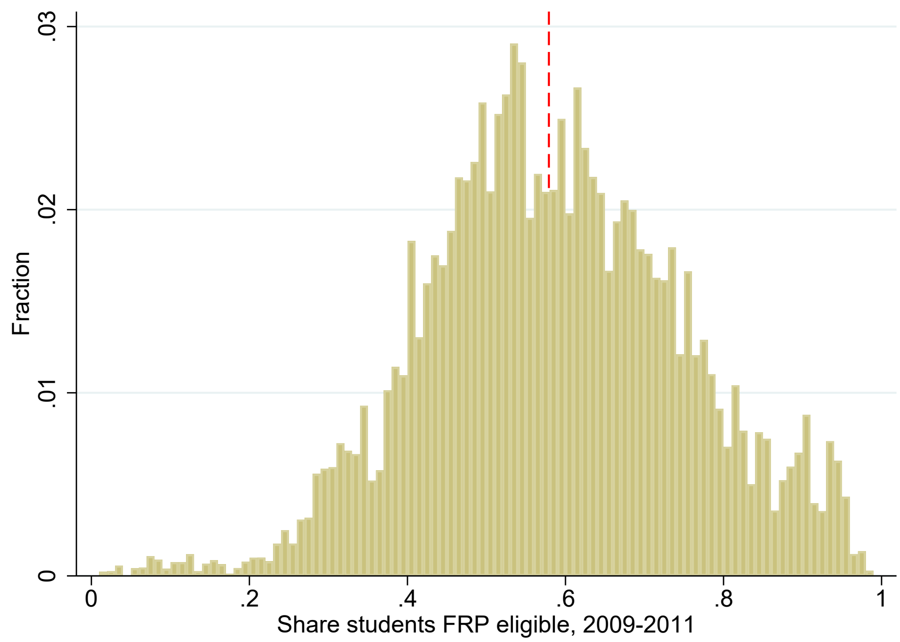
Notes: Figure shows the distribution of baseline (2009-2011) district characteristics by the year a district-grade first participated in CEP. "No CEP" are district-grades that did not adopt CEP by 2017. All dollars in constant 2017 dollars, adjusted for inflation with the CPI-U-RS. See text and data appendix for variable details.

Figure 6: Baseline District Income Assistance and Student Baseline Performance by CEP Adoption



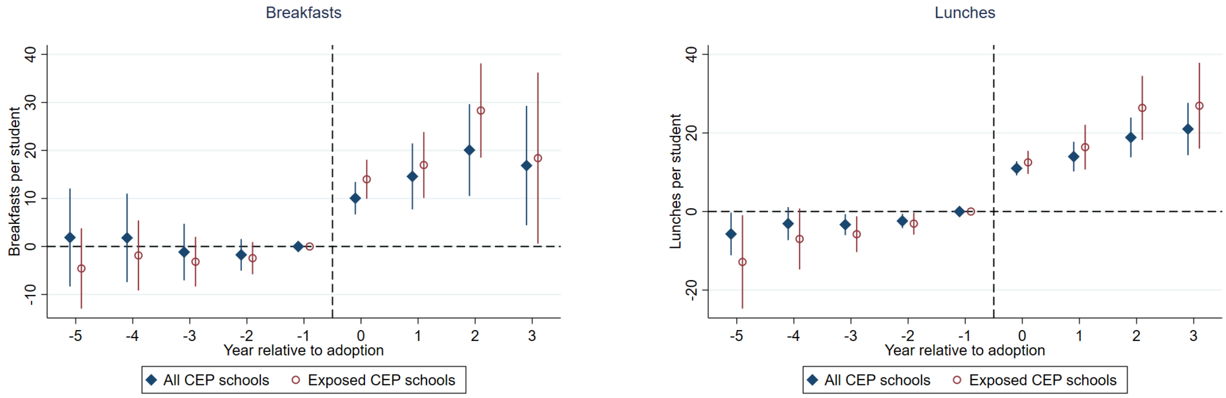
Notes: Figure shows the distribution of baseline (2009-2011) district characteristics by the year a district-grade first participated in CEP. "No CEP" are district-grades that did not adopt CEP by 2017. All dollars in constant 2017 dollars, adjusted for inflation with the CPI-U-RS. See text and data appendix for variable details.

Figure 7: Baseline FRP Eligibility, CEP-participating Districts



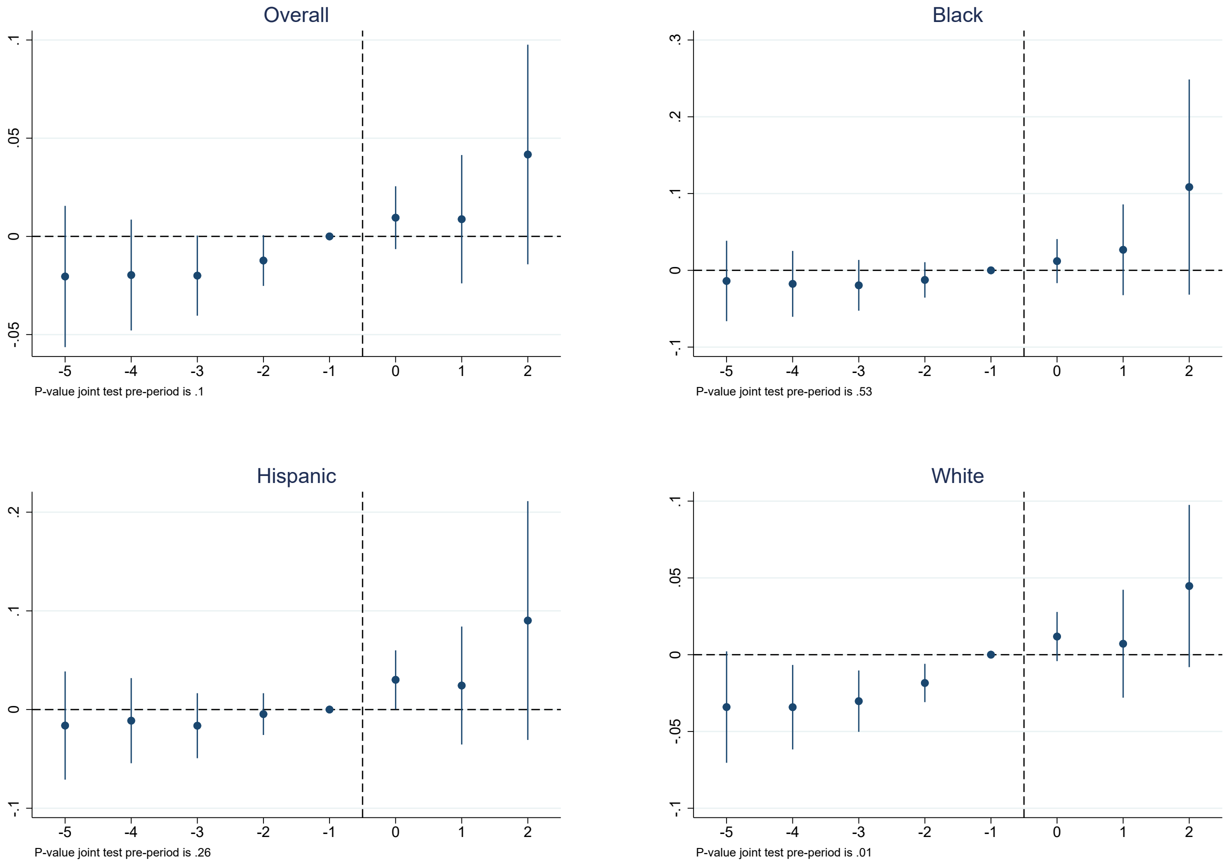
Histogram shows the distribution of the fraction of students in each CEP-participating district-grade who were eligible for free meals between 2009 and 2011. The dashed vertical line denotes the median eligibility rate among CEP-participating districts – 57.9 percent (the "exposed" cutpoint in the main analyses).

Figure 8: Event Study: Breakfast and Lunch Participation



Notes: Figure presents results from the (school-level) event study framework in Equation 2. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and school fixed effects. Bars denote 95 percent confidence intervals from robust standard errors clustered by district. Exposed subsample includes schools in districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). p-values for joint hypothesis that pre-period coefficients are different from zero are as follows: Breakfasts: 0.09 (all), 0.17 (exposed); Lunches: 0.02 (all), 0.01 (exposed).

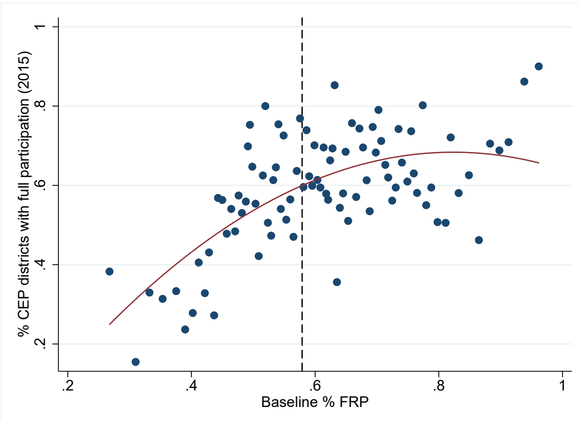
Figure 9: Math Performance Event Study



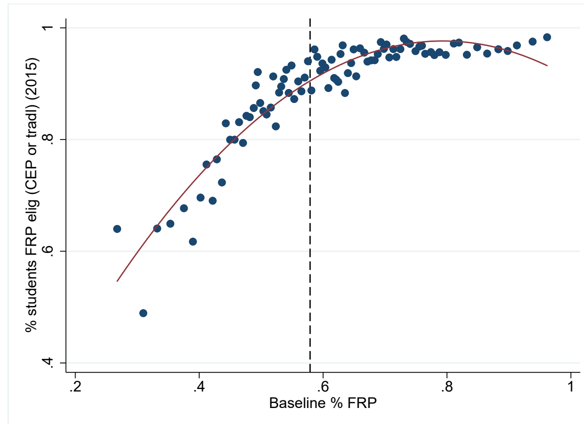
Notes: Figure presents results from the (district-level) event study framework in Equation 2. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and district fixed effects. Bars denote 95 percent confidence intervals from robust standard errors clustered by district. Sample includes districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Notes below each panel present p-values from the joint test that pre-treatment coefficients equal to zero.

Figure 10: Baseline Poverty and Free Meal Access

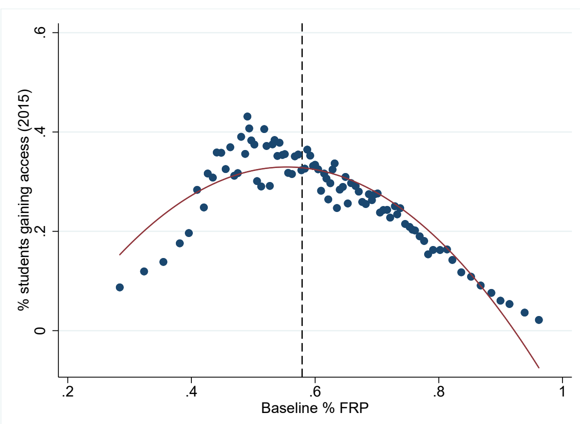
(a) Share of districts with full participation



(b) Share of students with access to free meals (either traditional or CEP)



(c) Share of students gaining access to free meals under CEP



Notes: Figure summarizes the relationship between various measures of CEP participation and access to free meals for all districts that adopted CEP by 2017. Dashed vertical line denotes the 57.9 percent threshold for the exposed subsample. Panel (a) presents the relationship between district baseline FRP eligibility and the probability of full district CEP participation; panel (b) shows the relationship between baseline FRP eligibility and the share of students with access to free meals after CEP was implemented (either students attending CEP schools or FRP students attending non-CEP schools); panel (c) shows the relationship between baseline eligibility and the fraction of students not previously eligible who gained access to free meals under CEP.

Table 1: Federal Reimbursement Rates for School Meals, 2017-2018

	Free	Reduced	Paid	Nutrition Quality
		Breakfast		
48 Contiguous States	\$1.66 (\$1.99)	\$1.36 (\$1.69)	\$0.29	
Alaska	\$2.66 (\$3.19)	\$2.36 (\$2.89)	\$0.43	
Hawaii	\$1.94 (\$2.32)	\$1.64 (\$2.02)	\$0.33	
		Lunch		
48 Contiguous States	\$3.23 (\$3.29)	\$2.83 (\$2.89)	\$0.31 (\$0.33)	\$0.06
Alaska	\$5.24 (\$5.26)	\$4.84 (\$4.86)	\$0.50 (\$0.52)	\$0.06
Hawaii	\$3.78 (\$3.80)	\$3.38 (\$3.40)	\$0.36 (\$0.38)	\$0.06

Notes: Source: USDA (2017). Left numbers show the base federal reimbursement rate; right numbers show the rate for high-poverty schools (for lunch, schools with at least 60 percent of students receiving free or reduced meals; for breakfast, schools with at least 40 percent of students receiving free or reduced meals). In addition, schools receive an additional 6 cents per lunch for serving fruits and vegetables.

Table 2: Baseline District Summary Statistics by Year of CEP Implementation

	(1)	(2)	(3)	(4)	(4)
	2012	2013	2014	2015	2016-7
Panel A: Baseline area characteristics					
% FRP	0.674	0.599	0.591	0.599	0.558
	(0.138)	(0.143)	(0.131)	(0.181)	(0.172)
Urban	0.158	0.191	0.142	0.170	0.131
	(0.365)	(0.393)	(0.350)	(0.376)	(0.337)
% college educated*	0.145	0.155	0.160	0.172	0.168
	(0.077)	(0.077)	(0.087)	(0.086)	(0.088)
Unemployment rate	0.077	0.064	0.057	0.060	0.055
	(0.034)	(0.029)	(0.019)	(0.027)	(0.026)
% single mother*	0.386	0.329	0.346	0.341	0.306
	(0.143)	(0.135)	(0.128)	(0.129)	(0.113)
Median household income ($1000s)*	40.490	44.410	45.340	46.200	49.710
	(12.890)	(10.990)	(11.400)	(13.730)	(14.970)
Gini coefficient*	0.432	0.416	0.414	0.411	0.397
	(0.060)	(0.043)	(0.051)	(0.048)	(0.052)
Child poverty rate*	0.319	0.272	0.259	0.264	0.237
	(0.127)	(0.092)	(0.090)	(0.110)	(0.104)
Per-capita income assistance ($1000s)	1.320	1.194	1.144	1.166	1.091
	(0.512)	(0.401)	(0.334)	(0.434)	(0.367)
Panel B: Baseline district characteristics					
% charter schools	0.016	0.048	0.045	0.034	0.026
	(0.048)	(0.108)	(0.091)	(0.077)	(0.076)
% black	0.306	0.205	0.299	0.241	0.186
	(0.343)	(0.277)	(0.277)	(0.306)	(0.253)
% Hispanic	0.075	0.063	0.086	0.201	0.171
	(0.133)	(0.122)	(0.132)	(0.287)	(0.260)
% special education	0.172	0.171	0.148	0.138	0.141
	(0.035)	(0.037)	(0.041)	(0.057)	(0.058)
# schools	16.970	21.010	20.500	17.480	14.810
	(41.870)	(64.490)	(39.510)	(33.790)	(47.550)
Student-teacher ratio	16.060	15.680	15.750	15.480	18.640
	(2.652)	(2.321)	(3.394)	(8.145)	(154.000)
Per-pupil expend ($1000s)	13.150	14.090	12.500	13.470	13.410
	(2.324)	(3.900)	(2.789)	(4.117)	(4.296)
School meal revenue ($1000s)	0.440	0.386	0.405	0.413	0.383
	(0.109)	(0.115)	(0.117)	(0.163)	(0.140)
% full CEP participation (2017)	0.710	0.601	0.649	0.605	0.611
	(0.417)	(0.465)	(0.447)	(0.461)	(0.459)

	(1)	(2)	(3)	(4)	(4)
	2012	2013	2014	2015	2016-7
Panel C: Baseline district performance					
Overall math	-0.382	-0.326	-0.253	-0.288	-0.242
	(0.346)	(0.326)	(0.303)	(0.368)	(0.358)
Hispanic math	-0.493	-0.585	-0.306	-0.366	-0.366
	(0.331)	(0.281)	(0.279)	(0.298)	(0.301)
White math	-0.207	-0.187	-0.059	-0.004	-0.017
	(0.326)	(0.310)	(0.326)	(0.332)	(0.335)
Black math	-0.721	-0.673	-0.582	-0.586	-0.556
	(0.282)	(0.282)	(0.276)	(0.297)	(0.300)
Overall reading	-0.278	-0.242	-0.188	-0.299	-0.238
	(0.320)	(0.300)	(0.280)	(0.342)	(0.354)
Hispanic reading	-0.427	-0.495	-0.332	-0.474	-0.456
	(0.338)	(0.270)	(0.285)	(0.265)	(0.303)
White reading	-0.108	-0.112	0.012	0.022	0.010
	(0.280)	(0.307)	(0.308)	(0.292)	(0.302)
Black reading	-0.599	-0.536	-0.508	-0.543	-0.510
	(0.274)	(0.261)	(0.246)	(0.261)	(0.275)
Observations	2162	4316	4310	26630	23314

Notes: Table shows baseline characteristics by year of CEP implementation for district-grades with any school participating in CEP between 2012 and 2017. Baseline defined as 2006-2010 for data available through the American Community Survey (denoted by "∗"), 2009-2011 for other sources. Column headers denote the first year a district-grade had any CEP participation. Unemployment rate from BLS LAUS, child poverty rates from Census Bureau SAIPE, per-capita income assistance from BEA REIS. Other area characteristics from the American Community Survey, and all district resources and performance measures from SEDA. All dollars in 2017 constant thousand dollars, adjusted for inflation with the CPI-U-RS. See text and data appendix for details.

Table 3: CEP and Change in Breakfasts and Lunches Served

	(1)	(2)	(3)	(4)	(5)	(6)
	Per-student breakfast		Per-student lunch		Log per student nutrition asst	
	All	Exposed	All	Exposed	All	Exposed
CEP	19.873***	19.794***	13.194***	12.140***	0.091***	0.093***
	(2.582)	(3.228)	(1.188)	(1.085)	(0.009)	(0.012)
Observations	18762	12077	20030	13193	128656	64212
Baseline DV mean	52.57	49.16	111.9	104.3	0.400	0.327
% change	0.378	0.403	0.118	0.116		
Level	School	School	School	School	District	District

Notes: Table presents regression results from unweighted school-level meal count data (columns 1-4) collected from state Department of Educations for six of the eleven states that adopted CEP before 2015: Georgia, Illinois, Kentucky, New York, Maryland, and West Virginia. Data availability varies by state, but spans 2009-2016. Columns 5 and 6 presents federal nutritional assistance to districts from the Annual Survey of School System Finances ($1,000s of 2017 dollars). All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, as well as year fixed effects. Columns 1-4 also include school fixed effects; columns 5 and 6 include district fixed effects. Odd numbered columns ("All") include all observations that adopted CEP between 2012 and 2017; even-numbered columns ("Exposed") restrict the sample to observations in districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 4: Effect of CEP on Academic Performance: Full Sample

	(1)	(2)	(3)	(4)	(5)
	Overall	Overall	Black	Hispanic	White
Panel A: Math performance					
CEP	0.002	0.004	-0.002	0.017	0.009
	(0.010)	(0.008)	(0.012)	(0.012)	(0.007)
Observations	65800	65800	30530	25258	51056
Baseline FRP	0.586	0.586	0.625	0.573	0.554
Baseline DV mean	-0.274	-0.274	-0.587	-0.376	-0.036
Panel B: Reading performance					
CEP	-0.012**	-0.006	-0.014	0.003	0.002
	(0.005)	(0.005)	(0.009)	(0.008)	(0.005)
Observations	68779	68779	31547	26207	52943
Baseline FRP	0.585	0.585	0.623	0.571	0.553
Baseline DV mean	-0.263	-0.263	-0.530	-0.460	0.001
Area and district controls		X	X	X	X
Sample	All	All	All	All	All

Notes: Table presents weighted least squares regression results from Equation 1 for all district-grade observations in which any school serving grade g participated in CEP by 2017. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects. "Area and district controls" include student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 5: Baseline District Summary Statistics by Baseline FRP Eligibility

	(1)	(2)
	Not exposed	Exposed
Panel A: Baseline area characteristics		
% FRP	0.718	0.453
	(0.111)	(0.110)
Urban	0.165	0.144
	(0.371)	(0.351)
% college educated*	0.148	0.187
	(0.074)	(0.093)
Unemployment rate	0.066	0.051
	(0.030)	(0.020)
% single mother*	0.386	0.272
	(0.134)	(0.085)
Median household income ($1000s)*	39.89	54.36
	(10.24)	(13.60)
Gini coefficient*	0.424	0.390
	(0.050)	(0.046)
Child poverty rate*	0.311	0.202
	(0.100)	(0.084)
Per-capita income assistance ($1000s)	1.291	0.996
	(0.437)	(0.311)
Panel B: Baseline district characteristics		
% charter schools	0.037	0.027
	(0.089)	(0.069)
% black	0.343	0.106
	(0.336)	(0.152)
% Hispanic	0.197	0.137
	(0.287)	(0.224)
% special education	0.140	0.147
	(0.062)	(0.047)
# schools	16.530	16.980
	(49.410)	(34.670)
Student-teacher ratio	17.890	15.610
	(135.200)	(4.595)
Per-pupil expend ($1000s)	13.730	13.100
	(4.143)	(3.937)
School meal revenue ($1000s)	0.474	0.328
	(0.143)	(0.111)
% full CEP participation (2017)	0.695	0.566
	(0.460)	(0.496)

	(1)	(2)
	Not exposed	Exposed
Panel C: Baseline district performance		
Overall math	-0.429	-0.121
	(0.346)	(0.300)
Hispanic math	-0.437	-0.315
	(0.308)	(0.285)
White math	-0.101	0.011
	(0.346)	(0.320)
Black math	-0.641	-0.502
	(0.294)	(0.281)
Overall reading	-0.421	-0.104
	(0.320)	(0.283)
Hispanic reading	-0.528	-0.391
	(0.283)	(0.269)
White reading	-0.067	0.051
	(0.314)	(0.281)
Black reading	-0.588	-0.441
	(0.257)	(0.254)
Observations	10433	10281

Notes: Table shows baseline characteristics by year of CEP implementation for district-grades with any school participating in CEP between 2012 and 2017. "Not exposed" describes participting district-grades with baseline FRP eligibility above 57.9 percent (the median among all CEP districts); "Exposed" described districts with baseline eligibility below 57.9 percent.Baseline defined as 2006-2010 for data available through the American Community Survey (denoted by "*"), 2009-2011 for other sources. Column headers denote the first year a district-grade had any CEP participation. Unemployment rate from BLS LAUS, child poverty rates from Census Bureau SAIPE, per-capita income assistance from BEA REIS. Other area characteristics from the American Community Survey, and all district resources and performance measures from SEDA. All dollars in 2017 constant thousand dollars, adjusted for inflation with the CPI-U-RS. See text and data appendix for details.

Table 6: Effect of CEP on Academic Performance: Exposed Districts Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Black	Hispanic	White	WB gap	WH gap
	Panel A: Math performance					
CEP	0.016*	0.028	0.034**	0.017*	0.020**	0.002
	(0.009)	(0.018)	(0.016)	(0.010)	(0.010)	(0.010)
Observations	32694	11658	12698	29325	11370	11465
Baseline FRP	0.454	0.457	0.438	0.458	0.457	0.449
Baseline DV mean	-0.121	-0.502	-0.315	0.0111	0.626	0.457
FRP gain	0.319	0.228	0.212	0.317	0.226	0.204
	Panel B: Reading performance					
CEP	0.007	0.016*	0.017*	0.008	0.004	0.005
	(0.006)	(0.010)	(0.010)	(0.007)	(0.010)	(0.009)
Observations	34344	12185	13256	30581	11894	11826
Baseline FRP	0.453	0.457	0.436	0.458	0.457	0.449
Baseline DV mean	-0.104	-0.441	-0.391	0.0509	0.592	0.532
FRP gain	0.319	0.230	0.209	0.317	0.228	0.200
Area and district controls	X	X	X	X	X	X
Sample	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed

Notes: Table presents weighted least squares regression results from Equation 1 for "exposed" district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. "FRP gain" is the share of students gaining access to free meals under CEP relative to the baseline (2009-2011) period. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 7: Effects of CEP on Academic Performance in Exposed Districts: By Grade and Race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall		Black		Hispanic		White	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Panel A: Elementary (grades 3-5)								
CEP	0.020*	0.007	0.055**	0.019*	0.050**	0.024	0.006	0.018
	(0.011)	(0.007)	(0.022)	(0.011)	(0.020)	(0.015)	(0.008)	(0.011)
Observations	17989	18419	6411	6551	7387	7457	16452	16089
Baseline FRP	0.456	0.456	0.455	0.455	0.437	0.437	0.460	0.460
Baseline DV mean	-0.082	-0.050	-0.458	-0.383	-0.288	-0.352	0.103	0.053
Panel B: Middle (grades 6-8)								
CEP	0.006	0.004	0.003	-0.017	0.008	0.006	0.008	0.010
	(0.008)	(0.010)	(0.014)	(0.016)	(0.012)	(0.017)	(0.009)	(0.011)
Observations	15925	14705	5634	5247	5799	5311	14129	13236
Baseline FRP	0.451	0.452	0.459	0.460	0.434	0.439	0.456	0.456
Baseline DV mean	-0.166	-0.167	-0.507	-0.552	-0.439	-0.349	-0.009	-0.038
Area and district controls	X	X	X	X	X	X	X	X

Notes: Table presents weighted least squares regression results from Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 8: Effect of CEP on District Resources, Exposed Districts

	(1)	(2)	(3)	(4)	(5)
	Log(fed revenue)	Log(all revenue - nutr asst)	Log(per-pupil expend)	Log(per-pupil instruct. expend)	Student-teacher ratio
CEP	0.020	0.005	0.009	-0.016**	-0.418**
	(0.012)	(0.006)	(0.011)	(0.006)	(0.212)
Observations	33356	33419	28747	28927	33419
Baseline FRP	0.453	0.453	0.453	0.453	0.453
Baseline DV mean ($1000s)	1.747	13.150	13.080	6.828	15.490
Student characteristics	X	X	X	X	X
Sample	Exposed	Exposed	Exposed	Exposed	Exposed

Notes: Table presents weighted least squares regression results from Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, percent of students attending a charter school, child poverty rates and county unemployment rates. All dollars in constant 2017 dollars, adjusted for inflation with the CPI-U-RS. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 9: Effect of CEP on School and District Student Composition

	(1)	(2)	(3)
	Black	Hispanic	White
	% district population		
CEP	0.001	-0.003***	0.001
	(0.001)	(0.001)	(0.001)
Observations	34454	34454	34453
Baseline DV mean	0.107	0.137	0.713
	% students CEP schools, partial participation districts		
CEP	0.009	0.003	0.003
	(0.006)	(0.004)	(0.002)
Observations	4115	4414	4872
Baseline DV mean	0.868	0.868	0.846
	Dissimilarity index, districts with multiple schools		
CEP	0.013**	0.001	0.008***
	(0.005)	(0.005)	(0.003)
Observations	26228	26228	26228
Baseline DV mean	0.280	0.256	0.249
Exposure	Exposed	Exposed	Exposed

Notes: Table presents unweighted regression results from Equation 1 for observations with a district baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. CEP equals one if any school serving grade g in district d participated in CEP by year t. Panel (a) indicates changes in district composition for each race/ethnic group r; panel (b) displays the fraction of students in racial/ethnic group r attending CEP schools; panel (c) indicates the district-grade racial/ethnic dissimilarity indices following initial CEP adoption. Panels (b) and (c) are limited to districts with multiple schools, panel (b) is limited to districts with partial CEP participation. All specifications include year, grade, and district fixed effects. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Table 10: Effects of CEP on Math Performance, Exposure Distribution

	(1)	(2)	(3)	(4)	(5)
Baseline FRP eligible	$\leq 40\%$	$\leq 50\%$	$\leq 60\%$	$\leq 70\%$	$\leq 80\%$
	Panel A: Overall performance				
CEP	0.009	0.017*	0.017*	0.012	0.009
	(0.018)	(0.010)	(0.009)	(0.009)	(0.009)
Observations	7654	20194	35798	49444	58815
Average baseline FRP	0.315	0.402	0.466	0.516	0.552
Baseline DV mean	-0.024	-0.079	-0.133	-0.187	-0.231
	Panel B: Black performance				
CEP	0.044	0.026	0.026	0.010	0.007
	(0.038)	(0.022)	(0.017)	(0.015)	(0.012)
Observations	2728	7282	13021	20188	26192
Average baseline FRP	0.326	0.407	0.471	0.535	0.583
Baseline DV mean	-0.451	-0.476	-0.511	-0.540	-0.564
	Panel C: Hispanic performance				
CEP	0.021	0.036*	0.031**	0.017	0.019
	(0.026)	(0.020)	(0.015)	(0.013)	(0.013)
Observations	3746	8194	13920	19226	23109
Average baseline FRP	0.299	0.383	0.452	0.505	0.545
Baseline DV mean	-0.284	-0.285	-0.321	-0.341	-0.362
	Panel D: White performance				
CEP	0.012	0.017	0.018*	0.011	0.010
	(0.020)	(0.011)	(0.009)	(0.008)	(0.007)
Observations	6456	18191	32039	43066	49214
Average baseline FRP	0.326	0.409	0.469	0.515	0.543
Baseline DV mean	0.122	0.042	0.006	-0.015	-0.031
Percentile baseline FRP distribution	11.700	31.000	54.800	75.500	89.600

Notes: Table presents weighted least squares regression results from Equation 1 for all district-grade observations in which any school serving grade g participated in CEP by 2017 based on the baseline (2009-2011) share of students FRP eligible under the traditional formula. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. "Average baseline FRP" indicates average baseline (2009-2011) eligibility rates. "Percentile baseline FRP distribution" displays the share of districts with baseline eligibility $\leq x\%$. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

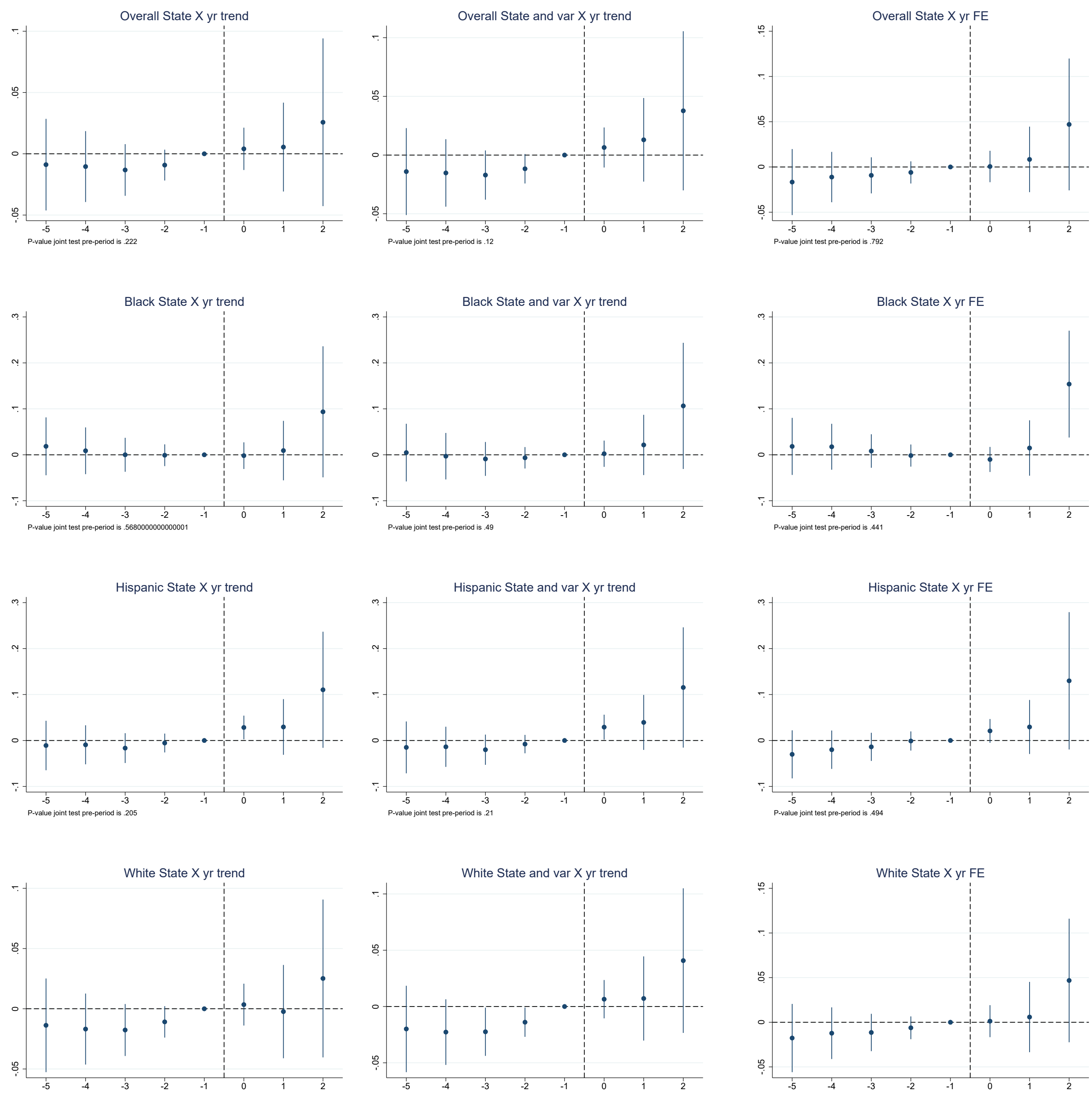
Table 11: Heterogeneous Effects of CEP on Math Performance, High Exposure Sample

	(1)	(2)	(3)	(4)
	High cost of living	Urban	% SNAP	Per-capita Income asst
	Panel A: Overall			
CEP	0.035***	0.016*	0.019*	0.019*
	(0.011)	(0.010)	(0.010)	(0.011)
CEP X char	-0.043***	-0.001	-0.013	-0.010
	(0.013)	(0.013)	(0.014)	(0.013)
Observations	32694	32694	32694	32694
p value: CEP + CEP X char	0.480	0.265	0.631	0.436
	Panel B: Black			
CEP	0.063**	0.033*	0.034*	0.035*
	(0.027)	(0.020)	(0.021)	(0.021)
CEP X char	-0.062**	-0.013	-0.039	-0.035
	(0.029)	(0.021)	(0.031)	(0.028)
Observations	11658	11658	11658	11658
p value: CEP + CEP X char	0.925	0.353	0.846	0.999
	Panel C: Hispanic			
CEP	0.054**	0.037**	0.028	0.033*
	(0.024)	(0.017)	(0.017)	(0.017)
CEP X char	-0.026	-0.007	0.031*	0.002
	(0.023)	(0.017)	(0.018)	(0.024)
Observations	12698	12698	12698	12698
p value: CEP + CEP X char	0.101	0.142	0.004	0.163
	Panel D: White			
CEP	0.032***	0.016	0.021*	0.022*
	(0.011)	(0.010)	(0.011)	(0.011)
CEP X char	-0.038***	0.009	-0.015	-0.014
	(0.014)	(0.015)	(0.017)	(0.015)
Observations	29325	29325	29325	29325
p value: CEP + CEP X char	0.629	0.118	0.721	0.587
Area and district controls	X	X	X	X
char $\geq$	91.600		0.271	1.082

Notes: Table presents weighted least squares regression results from Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. Race/ethnic proficiency scores available for cells with at least 20 students. CEP equals one if any school serving grade g in district d participated in CEP by year t. $CEPXchar$ equals one for districts with the baseline characteristic provided in the column header above the median. $char \geq$ displays the cutpoint for the interaction term (e.g.: districts with a regional purchasing power of at least 91.6 are considered "high-cost" areas). All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

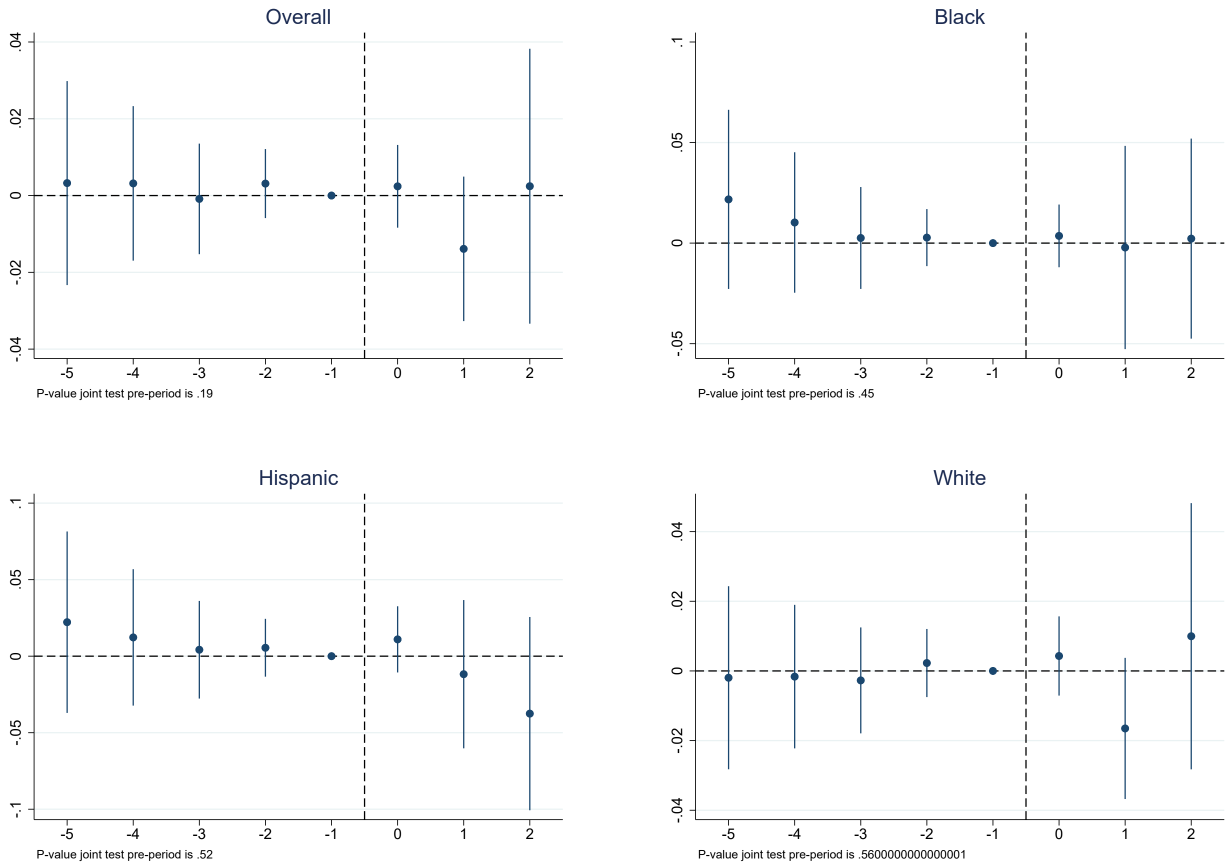
7 Appendix

Appendix Figure 1: Math Performance Event Study, Robustness, Exposed Districts



Notes: Figure presents results from event study framework in Equation 2. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and district fixed effects. Left and center panels include state linear trends, center panel also includes linear trends in baseline covariates. Right panel includes state-by-year fixed effects. Bars denote 95 percent confidence intervals from robust standard errors clustered by district. Sample includes districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Notes below each panel present p-values from the joint test that pre-treatment coefficients equal to zero.

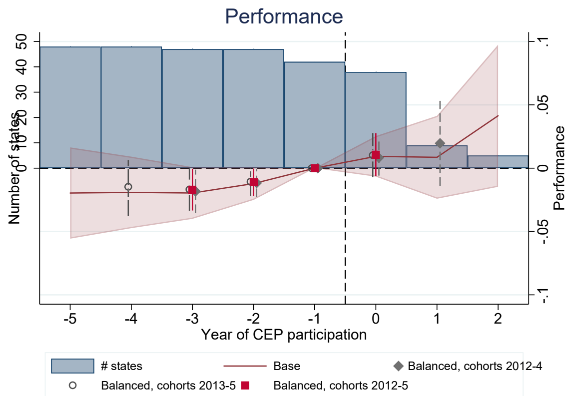
Appendix Figure 2: Reading Performance Event Study, Exposed Districts



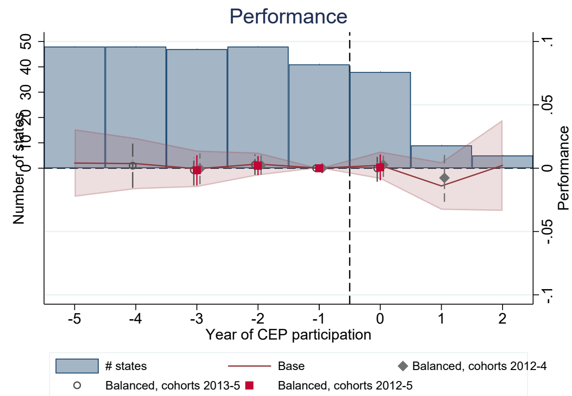
Notes: Figure presents results from the (district-level) event study framework in Equation 2. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and district fixed effects. Bars denote 95 percent confidence intervals from robust standard errors clustered by district. Sample includes districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Notes below each panel present p-values from the joint test that pre-treatment coefficients equal to zero.

Appendix Figure 3: Overall Performance: Balanced and Unbalanced Event Studies

(a) Math



(b) Reading



Notes: Figure summarizes the number of states contributing to each event year in the unbalanced panel (blue bars), and presents results from event study framework in Equation 2, with event years defined as year relative to CEP implementation for both unbalanced (maroon line) and three balanced subpanels. The gray diamonds show the balanced panel among districts that first adopted CEP between 2012 and 2014; the open gray circles show the 2013-2015 cohorts; and the bright red squares show the balanced event study for districts that adopted within the 2012 through 2015 period. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and district fixed effects. 95 percent confidence intervals from robust standard errors clustered by district. Sample includes districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts).

Appendix Figure 4: Math Performance: Drop Division, State, Grade

Notes: Figure plots coefficients and confidence intervals from the specifications in Table 6 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017, but dropping a single Census Division (blue), state (red), or grade (green). All omitted areas and grades are in ascending order (e.g.: the far-left point is Census Division 1, Alabama, or grade 3, the far-right point is Census Division 9, Wyoming, or grade 8. This figure indicates that results are not driven by the experiences of a single state or geographic area. Consistent with Table 7, math performance gains tend to be larger for younger grades. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Bars denote 95 percent confidence intervals from robust standard errors clustered by district.

Appendix Figure 5: Reading Performance: Drop Division, State, Grade

Notes: Figure plots coefficients and confidence intervals from the specifications in Table 6 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017, but dropping a single Census Division (blue), state (red), or grade (green). All omitted areas and grades are in ascending order (e.g.: the far-left point is Census Division 1, Alabama, or grade 3, the far-right point is Census Division 9, Wyoming, or grade 8. This figure indicates that results are not driven by the experiences of a single state or geographic area. Consistent with Table 7, math performance gains tend to be larger for younger grades. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Bars denote 95 percent confidence intervals from robust standard errors clustered by district.

Appendix Table 1: Effect of CEP on Meal Consumption: Parametric Event Study

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Breakfast				Lunch			
	All	All	Exposed	Exposed	All	All	Exposed	Exposed
Event year	5.982	4.981	2.601**	1.979	4.808	3.703	0.648	-4.487
	(3.460)	(3.419)	(0.970)	(5.410)	(2.966)	(2.813)	(1.229)	(2.635)
Post	9.983***	9.767***	13.330***	12.379***	10.245***	10.329***	10.433***	10.101***
	(1.719)	(1.608)	(1.838)	(1.749)	(0.851)	(0.828)	(1.036)	(1.100)
Event year X post	0.078	0.206	1.839	0.769	1.701	1.917	3.152***	3.231***
	(0.996)	(1.056)	(1.544)	(1.510)	(0.978)	(0.984)	(0.780)	(0.886)
StateXyear trends	X	X	X	X	X	X	X	X
Baseline var trends		X		X		X		X
Observations	14248	14248	6003	6003	14269	14269	6013	6013

Notes: Table presents unweighted results from meal count data collected from state Department of Educations for six of the eleven states that adopted CEP before 2015: Georgia, Illinois, Kentucky, New York, Maryland, and West Virginia. Data availability varies by state, but spans 2009-2016. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, as well as year and school fixed effects. Even-numbered columns also include state-specific linear trends and trends in baseline variables. Robust standard errors clustered by district. Columns (1-2) and (5-6) ("all") include all observations that adopted CEP between 2012 and 2017; columns (3-4) and (7-8) ("exposed") restrict the sample to observations in districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). $\beta_{ey} = \beta_{ey*post}$ presents p-value from a hypothesis test that the pre-CEP linear trends equals the trend after CEP adoption. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 2: Effect of CEP on Meal Consumption: Linear Trends by State and Baseline Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	School per-student breakfast		School per-student lunch		Log per student nutrit asst	
	All	Exposed	All	Exposed	All	Exposed
CEP	12.102***	12.520***	12.371***	12.129***	0.074***	0.082***
	(2.167)	(2.754)	(1.259)	(1.415)	(0.009)	(0.012)
Observations	18762	12077	20030	13193	128145	64105
Basline DV mean	52.57	49.16	111.9	104.3	0.400	0.327
Pct change	0.230	0.255	0.111	0.116		
StateXyear trends	X	X	X	X	X	X
Baseline trends	X	X	X	X	X	X
Level	School	School	School	School	District	District

Notes: Table presents unweighted results from estimating Equation 1 at the school level (columns (1) through (4)) with meal count data collected from state Department of Educations for six of the eleven states that adopted CEP before 2015: Georgia, Illinois, Kentucky, New York, Maryland, and West Virginia. Data availability varies by state, but spans 2009-2016. Columns 5 and 6 presents federal nutritional assistance dollars, reported in the Annual Survey of School System Finances. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, and measures of racial/ethnic segregation, as well as year and school fixed effects. Robust standard errors clustered by district. Odd-numbered columns ("all") include all observations that adopted CEP between 2012 and 2017; even-numbered columns ("exposed") restrict the sample to observations in districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 3: Federal Nutritional Assistance ($1,000s) and Overall Student Performance

	(1) Math	(2) Math	(3) Reading	(4) Reading
Per-student fed nutr. asst	0.163	0.512*	-0.142	0.235
	(0.212)	(0.296)	(0.142)	(0.206)
Observations	59465	31423	62174	32968
Sample	All	Exposed	All	Exposed
Baseline DV mean (level)	-.247	-.118	-.232	-.095
Change in nutritional asst	0.100	0.094	0.099	0.093
F stat 1st stage	184.856	106.867	156.994	105.841

Notes: Table presents 2SLS regression results where the change in per-student federal nutritional assistance is instrumented by CEP participation. "Exposed" districts are district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. "All" districts include all district-grade observations that participated in CEP at any point by 2017. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, percent of students attending a charter school, child poverty rates and county unemployment rates. All specifications are weighted least squares, with weights equal to the squared inverse of the standard error of the district-grade performance metric. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 4: Predicted Performance from Changes in Racial/Ethnic Composition

	(1)	(2)	(3)	(4)
	$\widehat{Overall}$	$\widehat{Black}$	$\widehat{Hispanic}$	$\widehat{White}$
	Panel A: Math performance			
CEP	-0.001	-0.002**	0.001	0.001
	(0.001)	(0.001)	(0.003)	(0.001)
Observations	32694	11658	12698	29325
Baseline FRP	0.454	0.457	0.438	0.458
Baseline DV mean	0.0773	-0.418	-0.225	0.214
	Panel B: Reading performance			
CEP	-0.000	-0.002**	-0.001	0.002
	(0.001)	(0.001)	(0.002)	(0.001)
Observations	34344	12185	13256	30581
Baseline FRP	0.453	0.457	0.436	0.458
Baseline DV mean	0.0745	-0.421	-0.226	0.213
Area and district controls	X	X	X	X
Sample	Exposed	Exposed	Exposed	Exposed

Notes: Table presents weighted least squares regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Dependent variable is defined as predicted values from a regression interacting each grade with the share of students of each racial/ethnic group in a district and CEP schools within a district, as well as the dissimilarity index for each racial/ethnic group. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 5: Effect of CEP on Math Performance: High-Exposure Districts Sample, Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A: All				
CEP	0.015	0.011	0.008	0.013	0.002	0.016*	0.018**
	(0.009)	(0.009)	(0.007)	(0.009)	(0.010)	(0.008)	(0.008)
Observations	32694	32607	26301	32694	32694	32694	32645
Baseline FRP	0.454	0.454	0.455	0.454	0.454	0.454	0.454
Baseline DV mean	-0.121	-0.121	-0.116	-0.121	-0.121	-0.121	-0.121
			Panel B: Black				
CEP	0.024	0.011	0.010	0.022	-0.002	0.001	0.006
	(0.018)	(0.016)	(0.014)	(0.020)	(0.021)	(0.014)	(0.014)
Observations	11658	11658	8996	11658	11658	11658	11658
Baseline FRP	0.457	0.457	0.459	0.457	0.457	0.457	0.457
Baseline DV mean	-0.502	-0.502	-0.500	-0.502	-0.502	-0.502	-0.502
			Panel C: Hispanic				
CEP	0.031**	0.029**	0.029***	0.027*	0.026	0.025*	0.029**
	(0.016)	(0.014)	(0.011)	(0.016)	(0.018)	(0.013)	(0.013)
Observations	12698	12679	9582	12698	12698	12698	12679
Baseline FRP	0.438	0.438	0.437	0.438	0.438	0.438	0.438
Baseline DV mean	-0.315	-0.315	-0.305	-0.315	-0.315	-0.315	-0.315
			Panel D: White				
CEP	0.017*	0.010	0.006	0.016*	-0.001	0.019**	0.021**
	(0.010)	(0.009)	(0.008)	(0.010)	(0.011)	(0.009)	(0.009)
Observations	29325	29272	23324	29325	29325	29325	29293
Baseline FRP	0.458	0.458	0.459	0.458	0.458	0.458	0.458
Baseline DV mean	0.011	0.011	0.015	0.011	0.011	0.011	0.011
Treatment defn	Binary	Binary	Binary	Binary	Pct	Binary	Binary
Resource variables	X						
State X year trends		X	X				
Baseline trends		X	X				
Lagged performance			X				
State X year FE				X			
Weights	WLS	WLS	WLS	WLS	WLS	District	Log enroll

Notes: Table presents regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. "Resource variables" include per-pupil total and instructional expenditures; "baseline trends" includes linear trends for baseline values of all control variables. Columns (1-4) present weighted least squares regressions with additional controls; column (5) presents unweighted results; column (6) weights each observation by the log number of students in each racial/ethnic group between 2009 and 2011. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 6: Effect of CEP on Reading Performance: High-Exposure Districts Sample, Alternative Specifications

	(1)	(2)	(3)	(4)		(6)	(7)
			Panel A: All				
CEP	0.007	0.005	0.001	-0.001	-0.002	0.007	0.008
	(0.006)	(0.005)	(0.004)	(0.005)	(0.007)	(0.007)	(0.006)
Observations	34344	34250	28329	34344	34344	34344	34291
Baseline FRP	0.453	0.453	0.454	0.453	0.453	0.453	0.453
Baseline DV mean	-0.104	-0.104	-0.101	-0.104	-0.104	-0.104	-0.104
			Panel B: Black				
CEP	0.014	0.008	0.001	0.006	-0.019	-0.007	-0.002
	(0.009)	(0.007)	(0.007)	(0.009)	(0.017)	(0.011)	(0.011)
Observations	12185	12185	9504	12185	12185	12185	12185
Baseline FRP	0.457	0.457	0.458	0.457	0.457	0.457	0.457
Baseline DV mean	-0.441	-0.441	-0.441	-0.441	-0.441	-0.441	-0.441
			Panel C: Hispanic				
CEP	0.016	0.018**	0.005	0.011	0.014	-0.000	0.006
	(0.010)	(0.009)	(0.009)	(0.010)	(0.015)	(0.011)	(0.011)
Observations	13256	13234	10110	13256	13256	13256	13236
Baseline FRP	0.436	0.436	0.434	0.436	0.436	0.436	0.436
Baseline DV mean	-0.391	-0.391	-0.385	-0.391	-0.391	-0.391	-0.391
			Panel D: White				
CEP	0.007	0.004	-0.000	-0.001	-0.004	0.010	0.011
	(0.007)	(0.006)	(0.005)	(0.006)	(0.008)	(0.007)	(0.007)
Observations	30581	30530	24789	30581	30581	30581	30550
Baseline FRP	0.458	0.458	0.459	0.458	0.458	0.458	0.458
Baseline DV mean	0.051	0.051	0.054	0.051	0.051	0.051	0.051
Treatment defn	Binary	Binary	Binary	Binary	Pct	Binary	Binary
Resource variables	X						
State X year trends		X	X				
Baseline trends		X	X				
Lagged performance			X				
State X year FE				X			
Weights	WLS	WLS	WLS	WLS	WLS	District	Log enroll

Notes: Table presents regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. Race/ethnic proficiency scores available for cells with at least 20 students. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. "Resource variables" include per-pupil total and instructional expenditures; "baseline trends" includes linear trends for baseline values of all control variables. Columns (1-4) present weighted least squares regressions with additional controls; column (5) presents unweighted results; column (6) weights each observation by the log number of students in each racial/ethnic group between 2009 and 2011. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 7: Effect of CEP on Math Performance: High-Exposure Districts Sample, Alternative Samples

	(1)	(2)	(3)
Panel A: All			
CEP	0.014	0.015	0.005
	(0.016)	(0.011)	(0.016)
Observations	14835	22162	12948
Baseline FRP	0.468	0.458	0.455
Baseline DV mean	-0.168	-0.098	-0.126
Panel B: Black			
CEP	-0.024	0.021	0.017
	(0.035)	(0.023)	(0.031)
Observations	2734	6636	5228
Baseline FRP	0.502	0.465	0.455
Baseline DV mean	-0.546	-0.487	-0.495
Panel C: Hispanic			
CEP	0.007	0.031	0.010
	(0.037)	(0.019)	(0.031)
Observations	2956	6552	6146
Baseline FRP	0.453	0.433	0.438
Baseline DV mean	-0.365	-0.273	-0.289
Panel D: White			
CEP	0.021	0.017	0.001
	(0.017)	(0.012)	(0.017)
Observations	12954	18949	11350
Baseline FRP	0.474	0.462	0.462
Baseline DV mean	-0.074	0.040	0.024
Sample	Full dist	Balanced	Adopt 1st yr

Notes: Table presents weighted least squares regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. Race/ethnic proficiency scores available for cells with at least 20 students. ‘All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Column (1) restricts to district-grade observations where every school serving grade g participates in CEP upon CEP adoption. Column (2) limits the sample to district-grade observations with a valid performance score each year. Column (3) limits the sample to districts that participated in CEP the first year their state became eligible. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 8: Effect of CEP on Reading Performance: High-Exposure Districts Sample, Alternative Samples

	(1)	(2)	(3)
	Panel A: All		
CEP	0.021*	-0.000	-0.021***
	(0.013)	(0.006)	(0.008)
Observations	15373	26999	13720
Baseline FRP	0.467	0.457	0.454
Baseline DV mean	-0.144	-0.096	-0.142
	Panel B: Black		
CEP	-0.007	0.005	0.003
	(0.027)	(0.010)	(0.012)
Observations	2782	7504	5397
Baseline FRP	0.501	0.462	0.455
Baseline DV mean	-0.472	-0.445	-0.462
	Panel C: Hispanic		
CEP	0.004	0.006	0.002
	(0.032)	(0.010)	(0.011)
Observations	3037	7028	6432
Baseline FRP	0.448	0.427	0.435
Baseline DV mean	-0.477	-0.388	-0.413
	Panel D: White		
CEP	0.024*	0.001	-0.028***
	(0.014)	(0.007)	(0.009)
Observations	13295	22302	11861
Baseline FRP	0.474	0.462	0.462
Baseline DV mean	-0.025	0.053	0.046
Sample	Full dist	Balanced	Adopt 1st yr
	participation	panel	eligibility

Notes: Table presents weighted least squares regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017; treatment districts are districts in which at least one school adopts CEP by 2015. Race/ethnic proficiency scores available for cells with at least 20 students. 'All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Column (1) restricts to district-grade observations where every school serving grade g participates in CEP upon CEP adoption. Column (2) limits the sample to district-grade observations with a valid performance score each year. Column (3) limits the sample to districts that participated in CEP the first year their state became eligible. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Appendix Table 9: Effects of CEP on Reading Performance, Exposure Distribution

	(1)	(2)	(3)	(4)	(5)
Baseline FRP eligible	$\leq 40\%$	$\leq 50\%$	$\leq 60\%$	$\leq 70\%$	$\leq 80\%$
	Panel A: Overall performance				
CEP	0.012	0.006	0.006	0.001	-0.003
	(0.013)	(0.007)	(0.006)	(0.006)	(0.005)
Observations	8054	21276	37599	51906	61616
Average baseline FRP	0.312	0.401	0.465	0.515	0.551
Baseline DV mean	-0.034	-0.065	-0.116	-0.175	-0.220
	Panel B: Black performance				
CEP	0.030	0.015	0.010	-0.003	-0.006
	(0.028)	(0.012)	(0.009)	(0.010)	(0.009)
Observations	2834	7631	13590	21013	27147
Average baseline FRP	0.325	0.407	0.471	0.534	0.582
Baseline DV mean	-0.383	-0.415	-0.450	-0.488	-0.510
	Panel C: Hispanic performance				
CEP	0.033	0.016	0.013	0.003	0.004
	(0.023)	(0.013)	(0.010)	(0.009)	(0.008)
Observations	3976	8605	14499	19961	23970
Average baseline FRP	0.294	0.380	0.449	0.503	0.543
Baseline DV mean	-0.358	-0.363	-0.394	-0.422	-0.445
	Panel D: White performance				
CEP	0.011	0.004	0.008	0.004	0.002
	(0.014)	(0.008)	(0.006)	(0.006)	(0.005)
Observations	6703	19017	33385	44762	51077
Average baseline FRP	0.326	0.410	0.469	0.514	0.542
Baseline DV mean	0.148	0.079	0.046	0.023	0.006
Percentile baseline FRP distribution	11.700	31.000	54.800	75.500	89.600

Notes: Table presents weighted least squares regression results from Equation 1 for all district-grade observations in which any school serving grade g participated in CEP by 2017 based on the baseline (2009-2011) share of students FRP eligible under the traditional formula. CEP equals one if any school serving grade g in district d participated in CEP by year t. Race/ethnic proficiency scores available for cells with at least 20 students. "Average baseline FRP" indicates average baseline (2009-2011) eligibility rates. "Percentile baseline FRP distribution" displays the share of districts with baseline eligibility $\leq x\%$. All specifications include district, grade, and year fixed effects, as well as student racial/ethnic composition and segregation, student-teacher ratios, percent of students attending a charter school, child poverty rates and county unemployment rates. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Data Appendix

CEP participation data

School-level CEP participation information comes from state educational agencies for the 2012 through 2014 pilot years. Data for the 2015 through 2017 academic years are provided for most states in the Common Core of Data, and for the remaining states by FRAC and CBPP. States with CEP participation in the CCD are identified by a unique NCES-issued district identifier. For states with incomplete CEP participation in the CCD (Illinois, West Virginia, the District of Columbia, North Carolina, Utah, Vermont, and Wyoming), I standardize all school and district names and match to the Common Core of Data (CCD) based on state, district, and school name. This procedure matches approximately 93 percent for all public elementary and middle schools.[25]

The CCD also provides grade-level enrollment information necessary to collapse the school-level CEP participation data to the grade-district level. The main analyses define a binary treatment variable equal to one if any school in district d serving grade g at time t had implemented CEP. For example, if in 2015, a district has a CEP school with some third graders, but no eighth graders, attending a school offering CEP, third graders are in the district are considered "treated," but eighth graders are considered "untreated." In practice, conditional on having *any* district participation in grade 3-8, 72 percent of districts have participation at each grade level, with slightly higher participation rates in younger grades. In robustness checks (Appendix Tables 5 and 6, I calculate a continuous measure of CEP participation, measured as the fraction of students in district d and grade g attending a CEP school in year t, using school-level enrollment data from the CCD.

Meal count data

No existing data set provides a consistent measure of school meal receipt across states or over time. Schools and states report this information to USDA for federal reimbursement;

[25] An earlier version of this paper matched schools using fuzzy string match to minimize the Levenshtein distance. These approaches yield almost identical empirical results; the current approach is more accurate.

however, the recorded meal count measures and the duration of maintained records varies by state. Between September 2016 and May 2017, research assistants contacted state staff in each of the eleven pilot states. Of these states, six provided school-level information on the number of breakfasts and lunches served in each school, and three provided this information by payment status. Following the string matching procedure described above, I match school-level meal receipt data to school-level CEP participation and demographic information from the CCD. In order to construct a comparable measure of meal consumption across states, I calculate the number of breakfasts and lunches per student, where student enrollment is provided in the CCD.

District performance data

I merge the district-level CEP data to district-grade-year performance data from the Stanford Education Data Archive (SEDA, version 2.1) using NCES district-level identifiers, with charter schools placed in the district in which they are geographically located. The SEDA data is a novel dataset that is unique in its ability to compare achievement across states and over time at the substate (e.g.: district) level by combining information from the state-level National Assessment of Educational Progress (NAEP) results and restricted-use state-proficiency data. Reardon et al. (2018) detail the data construction, variable definitions, and underlying assumptions. I highlight several key points here regarding measurement and sample restrictions, and discuss the feasibility of alternative performance measures.

Performance measurement:

As required by federal legislation, each state administers reading and math examinations to every student in grades 3-8. The number of students scoring "proficient" are reported at the school level. Each state has its own "proficiency" standard and these definitions vary both across states and over time.[26]

[26] A substantial change occurred in the 2014-2015 school year, when the 34 states with waivers from the No Child Left Behind proficiency standards were required to give "high-quality assessments" of "college- and career-ready standards." These examinations (e.g. PARCC and Smarter Balance) tended to have more

The SEDA framework first applies a heteroskedastic ordered probit (HETOP) model (homoskedastic ordered probit for states with only two proficiency categories) to the proficiency categories reported to the Department of Education to estimate the mean and standard deviation for each state-subject-grade-year at the district level, as well as the corresponding standard errors. Intuitively, this step transforms categorical proficiency measures to a continuous measure. These means and standard deviations are then standardized by the state-subject-grade-year distribution to have a mean of zero and a standard deviation of one. Scores for each race/ethnic subgroup are based on the overall cutpoints so that the performance scale for each subgroup matches the overall distribution.

In order to facilitate comparisons across states and over time, the SEDA data adjusts these state-grade-year-subject estimates with information from state-level National Assessment of Educational Progress (NAEP) results. NAEP is an examination that is administered biennially to fourth and eighth graders in a sample of districts, and is designed to yield measures of math and reading achievement that are comparable across states and over time (see Reardon et al. (2017) for a detailed discussion). Each test wave, NAEP is designed to be nationally representative (and representative at the state level for math and reading), but only a sample of schools and students are chosen to participate, which limits its ability to track student performance at the school- or district-level over time.[27]

The SEDA framework takes the state-level NAEP data and first interpolates and extrapolates each state mean and standard deviation to years and grade levels not covered in the NAEP (e.g.: even-numbered years and grades 3 and 5-7). It then places the district-grade-subject-year continuous proficiency measures from the state-assessment data on the cross-state NAEP scale. A technical discussion is provided in Reardon et al. (2018) (Equations 6.2 and 6.3). At an intuitive level, districts that perform well on their state's assessment are placed on the SEDA scale high relative to their state's NAEP measure, and districts in states that score higher on the NAEP assessment also place higher in the SEDA performance

stringent proficiency requirements than the earlier state examinations Education Week (2014) and resulted in a sharp drop in the number of students achieving proficiency.

[27]District-level results are available for 21 of the largest districts that participate in the Trial Urban District Assessment (TUDA) pilot.

distribution.

Finally, the NAEP-proficiency estimates are standardized to the national student-level NAEP score distributions (for years and grades the NAEP was administered) or the interpolated/extrapolated NAEP score distributions (for observations the NAEP was not administered). All results in this paper use the "cohort standardized (cs)" scale, which provides a measure of performance that is comparable over time at the district-grade-subject level ((Reardon et al., 2018) Equation 7.1). Under this measure, treatment effects are provided in effect sizes.

Following Reardon et al. (2018), all models are estimated using weighted least squares, with weights inversely proportional to the estimated variance of the performance metric. These weights, detailed in Reardon et al. (2019) account for the NAEP interpolation and extrapolations, as well as other sources of linking error.

Excluded observations

Performance measures are not included for district-years that administer locally-selected examinations (this mainly affects middle schools in California, Virginia, and Texas in some years), district-grade-years with participation rates below 95 percent, instances where district-grade-year information was not reported to the Department of Education, cases of identified data errors in state proficiency data, or cases with estimated standard errors greater than the state-standardized metric. Finally, random noise is added to each estimate and district-subject-grade-year cells derived from fewer than 20 assessments are suppressed due to confidentiality issues. As this random noise introduces classical measurement error, it will slightly attenuate the reported results. Data Appendix Table 1 lists the state-subject-grades that are not available; Reardon et al. (2018) further describes the rationale for exclusion.

The final dataset has approximately 66,000 (math) to 69,000 (reading) district-year-grade observations for locations that participated in CEP through 2017, with subgroup sample sizes ranging from 25,000 (Hispanic math performance) to 53,000 (white reading performance). My main results focus on the 32,000-34,000 district-grade-year observations with relatively low district eligibility for the free meals program before CEP (those where less than 57.9 percent of students were eligible for free meals between 2009-2011).

In order to verify that CEP adoption is not associated with whether a valid score exists for each district-grade-subject-year, I create a balanced panel of district-grade-subjects and create a binary outcome variable equal to one if the district-grade-subject-subgroup-year is available in the SEDA data. Data Appendix Table 2 presents the results for math and reading and shows for overall, Hispanic, and white subgroups, there is no economically or statistically significant relationship between CEP implementation and the availability of SEDA data. For black performance, CEP adoption is associated with a 2 percentage point lower probability a district-subject-grade-year appears in the SEDA data. However, the results from the balanced sample of district-grades in Column (2) of Appendix Tables 7 and 8 indicate that differential attrition from the SEDA data is not driving the main results.

Data Appendix Table 1: Grade-subject-years not available in the SEDA data

	Math							Reading						
	2009	2010	2011	2012	2013	2014	2015	2009	2010	2011	2012	2013	2014	2015
Arkansas	8	8					8							
California	7-8	7-8	7-8	7-8	7-8	3-8							3-8	
Colorado	3-8	3-8	3-8					3-8	3-8	3-8				
Connecticut						3-8							3-8	
Delaware							8							8
Florida						3-8								
Idaho						3-8							3-8	
Kansas						3-8							3-8	
Maine							6-8							6-7
Maryland						3-4, 6-7							3-4, 6-7	
Missouri					8	8	8							
Montana						3-8	3-8						3-8	3-8
Nebraska	3-8	3-8						3-8						
Nevada						3-8	3-8						3-8	3-8
New Hampshire							8							8
New Jersey							3-8							3-8
New York						6-8	3-8						6-8	3-8
North Dakota							3-8							3-8
Ohio							8							
Oklahoma				8	8									
Oregon						3-8							3, 7-8	
Rhode Island							6-8							5-8
South Dakota						3-8								3-8
Tennessee							8							
Texas				7-8	7-8	7-8	7-8							
Utah	8	8	8	8	8									
Virginia	5-8	5-8	5-8	5-8	5-8	5-8	5-8							
Washington						3-8	3-8						3-8	3-8
West Virginia				3-8		3-7					3-8			
Wyoming		3-8			3-8	3, 7-8					3-8		3-8	

Source: Reardon et al. (2018) Table A1

Data Appendix Table 2: SEDA availability and CEP adoption

	(1)	(2)	(3)	(4)
	SEDA exists			
	Overall	Black	Hispanic	White
	Panel A: Math performance			
CEP	-0.002	-0.021***	0.001	-0.000
	(0.004)	(0.008)	(0.008)	(0.008)
Observations	37114	37114	37114	37114
Baseline FRP	0.454	0.454	0.454	0.454
Baseline DV mean	0.949	0.334	0.347	0.840
	Panel B: Reading performance			
CEP	0.001	-0.018**	0.006	0.011
	(0.004)	(0.008)	(0.009)	(0.008)
Observations	37114	37114	37114	37114
Baseline FRP	0.453	0.453	0.453	0.453
Baseline DV mean	0.967	0.342	0.356	0.853
Sample	Exposed	Exposed	Exposed	Exposed

Table presents weighted least squares regression results from the specification in Equation 1 for district-grade observations with a baseline FRP eligibility share below 57.9 percent (the baseline median among CEP districts) in which any school serving grade g participated in CEP by 2017. Dependent variable equals one if a measure of student performance exists for the district-grade-subject-year in the SEDA data. Robust standard errors clustered by district. See text and data appendix for details. *** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.10$.

Other performance measures

The SEDA data is only data source that provide substate measures of academic performance that are consistently available at the substate level and are comparable over time. Here I discuss other measures of student performance that are less well-suited to this analysis.

First, the NAEP data are not administered to the same sample of schools each wave, limiting the potential for these data to draw comparisons within a district over time.

Second, as described in Reardon et al. (2018), state proficiency examinations change over time, both in the assessment battery and the proficiency definition. Therefore, comparing outcomes across states using state proficiency data is problematic, even in models that include state and year fixed effects. Of particular concern for this analysis, the pilot and national implementation CEP period coincides with three important changes to state proficiency examinations:

1. 2010-2013: Under the original No Child Left Behind (NCLB) legislation, schools were required to have all students scoring "proficient" by the 2014 school year (Institute of Education Sciences, 2009). As illustrated in Data Appendix Figure 1, there is very little variation in the proficiency measure over this period – in 29 states, most students were proficient in more than 95 percent of schools.

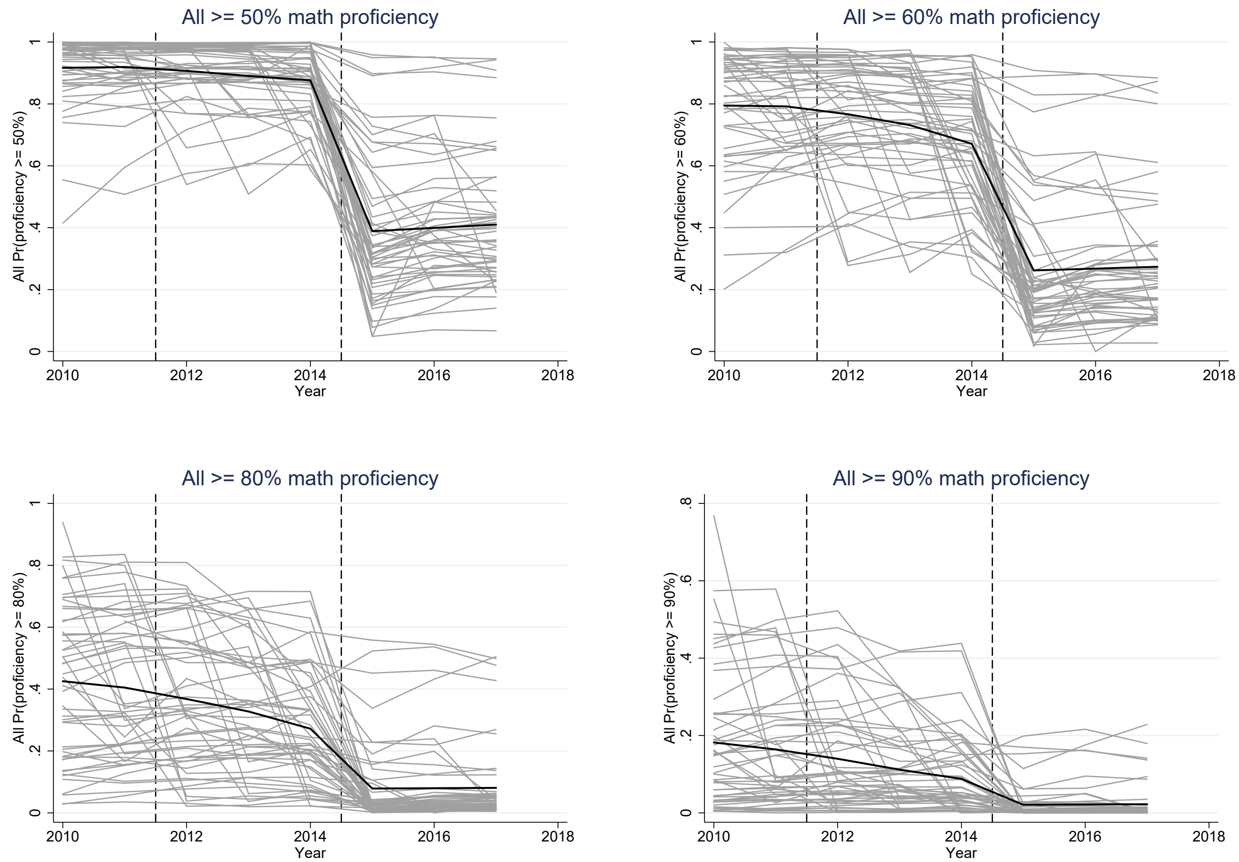
2. 2014: In the 2014 school year, 34 states received NCLB waivers, which provided exemptions from key elements of NCLB, including the proficiency requirements (US Department of Education, 2013)

3. 2015: States that received NCLB waivers are required to give "high-quality assessments" of "college- and career-ready standards." These examinations (e.g. PARCC and Smarter Balance) tended to have more stringent proficiency requirements than the earlier state examinations Education Week (2014). Accordingly, while there is within-state variation in the share of schools with proficiency marks upwards of 80 percent in the 2010-2013 period, there is little variation at these thresholds beginning in 2015 (probabilities of at least 80 percent students reaching proficiency falling close to 0,

particularly for Hispanic and black students). In 2017, in 45 states, less than 5 percent of schools had achieved 90 percent proficiency (Data Appendix Figure 1).

These limitations are particularly challenging over the 2009 through 2015 period as public-use data only provide proficiency shares within wide performance bins (e.g.: at least 50 percent of students proficient), making it difficult to discern modest changes in performance.

Data Appendix Figure 1: Fraction of Students Achieving Math Proficiency, by State



Notes: Figures show the fraction of schools in each state with at least $x\%$ of students in each race/ethnic group achieving the state proficiency measure. Each line corresponds to one state; thick black line corresponds to the national average. Dashed vertical lines indicate 2012 (first pilot year of CEP) and 2015 (first year all states became eligible). All data from Department of Education EdFacts. Proficiency measures are available for schools with at least 6 (x = 50%), 16 (x = 60, 80%), or 31 (90%) students taking the examination.

Additional control variables

The SEDA data include a rich set of covariates for the geographic district (e.g.: including information from charter schools located in separate administrative districts, but the same geographic district as public schools). As with the performance data, (Reardon et al., 2018) provide a comprehensive description. From the SEDA data, I include control variables for the share of students in a district black, Hispanic, special education, or English Language Learners, as well as the fraction of students attending a charter school and student-teacher ratios derived from the CCD. Summary statistics for baseline economic variables, including median household income, the share of female-headed households, and the Gini coefficient, derived from the 2006-2010 pooled American Community Survey.

In addition to these covariates, I merge data to each school district from several outside sources. The Census Bureau Small Area Income and Poverty Estimates (SAIPE) program provides child poverty rates; the Bureau of Labor Statistics Local Area Unemployment Statistics (LAUS) provides annual county unemployment rates. Data on district finances by level of government and type of revenue is provided by the LEA School District Finance Survey (F-33). Baseline rates of SNAP receipt are provided at the county level through USDA, and county-level per-capita income transfers are available through the Bureau of Economic Analysis Regional Economic Accounts (REIS) data. Finally, I explore heterogeneity by the local cost of living using the Bureau of Economic Analysis Regional Purchasing Parity (RPP) index.

I also augment the district CCD tabulations with several measures of *school* characteristics from the Common Core of Data. In particular, I estimate the fraction of black and Hispanic students attending CEP schools, and use these school-level counts to compute segregation measures for each race/ethnic group in a district-grade. The school district CCD data are also used to estimate the fraction of students gaining access to free meals under CEP relative to the traditional program (school-level data is necessary for this calculation for districts with partial CEP participation), as well as the continuous treatment measure presented in Appendix Tables 5 and 6.

Treatment definition

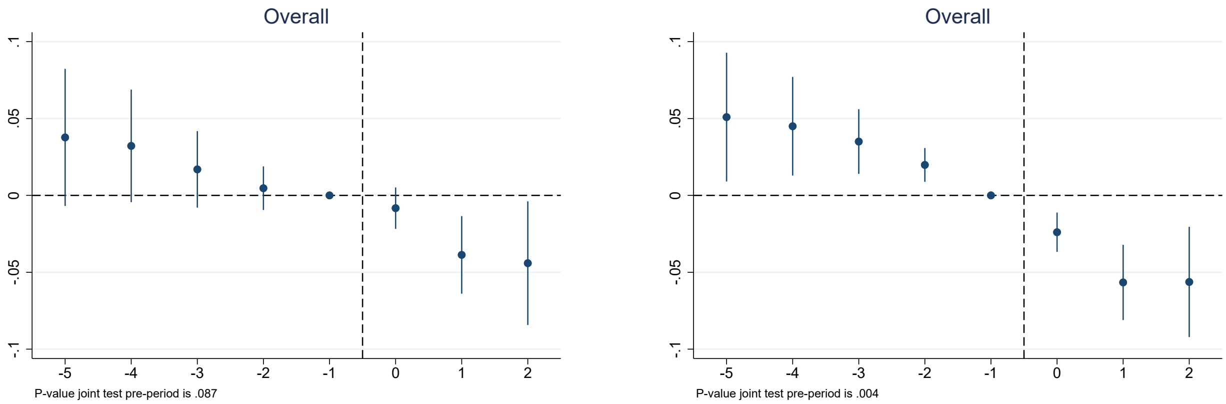
The analyses define CEP participation as the year in which a district-grade first participates in CEP and all subsequent years. Therefore, treatment depends on both the state in which a district is located and district-level decisions when to implement schoolwide free meals. In exploratory analyses, I have explored the feasibility of using the year each participating district became eligible for CEP (based on state) as the treatment variable. These analyses illustrate that using the timing of actual participation is better suited to evaluating the effect of CEP on student achievement than an eligibility-based treatment measure for two reasons.

First, calendar time and eligibility year are highly collinear: CEP was rolled out over a four-year period, with most schools (in 40 states) becoming eligible in 2015. Therefore, models using eligibility year as the treatment variable are unable to account for concurrent state policy changes that may affect both districts' (eventual) CEP participation and student achievement. In contrast, models that leverage actual implementation can account for these factors by augmenting the baseline specification with year-by-state fixed effects or state-specific linear trends. Appendix Figure 1 and Appendix Table 5 shows that math results are largely robust to these modifications. In contrast, reading results (Appendix Table 6) are more sensitive to the specification. Overall, it does not appear that CEP led to systematic changes in reading performance.

Second, it is possible that eligibility reflects selection at the state-level into treatment. If states were selected in part because of potential gains from CEP participation (or factors correlated with potential improvements), an eligibility-based treatment measure will not provide a biased measure of the causal effect of schoolwide free meals on student performance. Even with non-random selection of pilot states, however, the choice to participate in CEP is a school and district-level decision. The patterns shown in the eligibility-based event study plots in Data Appendix Figure 2 are consistent with negative selection of pilot states, although this is not conclusive evidence as it is difficult to disentangle selection from secular trends. In comparison, for black and Hispanic students, the timing of actual participation does not coincide with trends in student performance (Figure 9).

Data Appendix Figure 2: Performance: Academic Performance Event Study: Eligbility-defined Treatment, Exposed Districts

(a) Math performance (b) Reading performance



Notes: Figure presents results from event study framework in Equation 2 where event time is defined relative to the first year of CEP eligibility (2012 for districts in Michigan, Illinois, and Kentucky; 2013 for districts in the District of Columbia, New York, Ohio, and West Virginia; 2014 for districts in Florida, Georgia, Massachusetts, and Maryland; and 2015 for the remaining states). Outcome variable is overall math (panel (a)) or reading (panel (b)) performance subgroup results show similar patterns. All specifications include controls for student demographics, the fraction of charter schools in a district, child poverty and unemployment rates, measures of racial/ethnic segregation, year fixed effects, grade fixed effects, and district fixed effects. Bars denote 95 percent confidence intervals from robust standard errors clustered by district. Sample includes districts with a baseline FRP eligibility rate below 57.9 percent (the median among CEP-adopting districts). Notes below each panel present p-values from the joint test that pre-treatment coefficients equal to zero.