NATIONAL
CENTER *for* ANALYSIS *of* LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

CALDER

AIR

Using Free Meal and Direct Certification Data to Proxy for Student Disadvantage in the Era of the Community Eligibility Provision

Cory Koedel

Eric Parsons

WORKING PAPER No. 214-0119-1 • January 2019

Using Free Meal and Direct Certification Data to Proxy for Student Disadvantage in the Era of the Community Eligibility Provision

Cory Koedel
University of Missouri/CALDER

Eric Parsons
University of Missouri

Contents

Acknowledgments

We thank the Missouri Department of Elementary and Secondary Education for data access and Yang An and Jing Song for research assistance. This research was also supported by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see www.caldercenter.org/about-calder. All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders, the Missouri Department of Elementary and Secondary Education or the institutions to which the authors are affiliated.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any opinions, findings, and conclusions expressed in these papers are those of the authors and do not necessarily reflect the views of our funders.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street NW, Washington, DC 20007
202-403-5796 • www.caldercenter.org

Using Free Meal and Direct Certification Data to Proxy for Student Disadvantage in the Era of the Community Eligibility Provision

Cory Koedel, Eric Parsons
CALDER Working Paper No. 214-0119-1
January 2019

Abstract

Free and reduced-price meal (FRM) data are used ubiquitously to proxy for student disadvantage in education research and policy applications. The Community Eligibility Provision (CEP)—a recently-implemented, federally-administered program—allows schools serving low-income populations to identify all students as FRM-eligible regardless of individual circumstances. The substantive implications of this policy aside, we study its effect on FRM eligibility as a proxy for student disadvantage, and relatedly, we examine the viability of direct certification (DC) status as an alternative disadvantage measure. While there is some informational degradation in FRM data caused by the CEP, primarily with respect to capturing school-level disadvantage, we show that the impact of the CEP is generally modest. In the post-CEP era, DC and FRM data are similarly informative. Using both measures together can improve the identification of disadvantaged students in administrative data, but only marginally.

1. Introduction

The use of free and reduced-price meal (FRM) eligibility as a proxy for student disadvantage is ubiquitous in education research. Policymakers and regulators also rely on FRM data in their efforts to monitor and regulate educational outcomes and interventions, including the allocation of Title-I funding and for Title-I accountability (Camera, 2019; Riddle, 2015). It is common knowledge that FRM-eligibility is a noisy and coarse proxy for student poverty (Bass, 2010; Chingos, 2016; Harwell and LeBeau, 2010; Michelmore and Dynarksi, 2017); but while imperfect, it has been shown to be an effective indicator of disadvantage (Domina et al., 2018).

Research on the usefulness of FRM eligibility as a proxy for student disadvantage pre-dates the implementation of the Community Eligibility Provision (CEP), a federal program operated by the United States Department of Agriculture. The CEP allows all students in schools and districts serving low-income populations to receive free meals regardless of each student's individual circumstances. Setting aside the substantive effects of the CEP, which have been studied elsewhere (e.g., see Gordon and Ruffini, 2018), our focus is on how the CEP affects the proxy value of FRM eligibility as a measure of student disadvantage. We also explore the viability of direct certification (DC) status as an alternative measure that could be used in the post-CEP era in addition to, or in place of, FRM eligibility.

We assess CEP-induced change in the informational value of FRM data by documenting the ability of FRM data to predict key student outcomes—test scores and attendance. This approach builds on recent related work by Domina et al. (2018) and Michelmore and Dynarski (2017). It is motivated by a measurement error framework in which the CEP can be thought of as increasing measurement error in the FRM-eligibility indicator. The nature of measurement error induced by the CEP is complicated and its substantive importance is unclear *a priori*, prompting our empirical investigation.

Our analysis is based on administrative microdata from Missouri. We begin by using the coded FRM status of students before and after the CEP was introduced to predict student outcomes. The pre/post models are a useful starting point for thinking about the influence of the CEP, but inference is confounded by changes to contextual factors over time that coincide with its introduction (e.g., changes to economic conditions, testing instruments, etc.). In order to separate the effect of the CEP from other factors we rely on "pseudo-coded" scenarios in which we falsely code schools as CEP adopters prior to policy implementation. That is, we use pre-CEP data and overwrite student FRM eligibility as if the CEP were in place. This allows us to hold everything other than the FRM coding status of students constant in our comparisons.

In the initial pseudo-coding scenario, we look forward in the data and identify Missouri schools that adopted the CEP in the first year it was available. We then go back in time and pseudo-code these schools as CEP adopters prior to the policy and estimate our models using the pre-CEP, pseudo-coded data. By comparing the results to results from models that use the actual pre-CEP data and FRM coding, we can assess how CEP-induced changes to which students are coded as FRM-eligible affects the informational content of FRM eligibility, holding all else constant. We expand on this basic idea to include schools that adopted the CEP within the first three years of program availability, then obtain an upper bound effect of the CEP by pseudo-coding all CEP-eligible schools in Missouri regardless of their future adoption decisions.

We show that the informational degradation in FRM eligibility attributable to the CEP is not zero, but it is small. This is true particularly for FRM eligibility measured at the individual level; the information contained by schoolwide FRM eligibility is degraded somewhat more. There are two mechanisms that account for our generally modest findings in this regard. First, students whose coding status is changed by the CEP are not a random sample of students—they are already a disadvantaged group as evidenced by their attendance at high-poverty schools. While these students are "miscoded" in a technical sense because of the CEP, the substantive implications of the miscoding are modest. Second, and more importantly, we show that the number of students who experience a FRM status change due to the CEP—even in the extreme scenario where all eligible schools in Missouri adopt the CEP—is small. This result may be initially surprising but is easy to explain *ex post*. The reason is that schools eligible for the CEP already have high shares of FRM-eligible students (about 80 percent on average), so relatively few students switch status when a school adopts the CEP. At its maximum effect, we show that the CEP would increase the share of FRM-eligible students in the state of Missouri by just 5.3 percentage points, raising it from 51.2 percent to 56.5 percent.

After documenting the effect of the CEP on the informational content of FRM data, we examine the potential for direct certification (DC) status to replace or augment FRM data in efforts to identify disadvantaged students in the post-CEP era. DC data have been suggested as an alternative to FRM data in recent policy reports and in the popular press (Camera, 2019; Chingos, 2016; Greenberg, 2018). The results from our comparative analysis generally show that FRM and DC data are similarly predictive of student outcomes in the post-CEP era, although there is some heterogeneity in their relative predictive power across outcomes. Specifically, FRM data are somewhat more predictive of student test scores than DC data, while the reverse is true for student attendance.

These findings have several implications for research and policy applications. First, despite some degradation of information in FRM data due to the CEP, FRM eligibility continues to proxy for student disadvantage in the post-CEP era at roughly the same level of efficacy as in the pre-CEP era. Our analysis makes clear that DC data are a viable substitute for FRM data, but also shows that DC data do not offer a meaningful improvement. Somewhat surprisingly, the marginal value of using DC data to augment FRM data in an effort to better identify student disadvantage, while positive, is small.

2. The Community Eligibility Provision

The CEP is a program run by the United States Department of Agriculture (USDA). It allows high poverty schools and districts to provide free meals (breakfast and lunch) to all students without collecting individual household applications. School and district eligibility for the CEP is based on the fraction of students who are "directly certified," which means they participate in other means-tested programs such as the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), and the Food Distribution Program on Indian Reservations. Students can also be directly certified if they are classified as foster, migrant, homeless, or runaway.

Schools and districts choose whether to participate in the CEP, conditional on eligibility. Participating institutions are reimbursed for the free meals by the USDA using a kinked formula based on the share of DC students. The DC share must be at least 0.40 for baseline eligibility. The CEP reimburses schools and districts for the free meals at a rate of 1.6 times the DC share and once the DC share reaches 0.625, the reimbursement rate plateaus at 100 percent. A final notable feature of the program is that when a school or district is accepted, it can offer free meals and receive reimbursement for four years without the need to re-apply. Our data panel covers the first three years of CEP implementation in Missouri (see below)—therefore, schools that we observe implementing the CEP remain covered throughout the timeframe we study.

We leverage CEP program rules for portions of our analysis to identify CEP-eligible schools in Missouri. We define eligibility as meeting the 0.40 DC share, which is the broadest definition. It is useful to map the poverty level of eligible schools by this definition to the FRM share. To do this we use data from 2014, which is the year before any Missouri schools adopted the CEP, thus preserving the informational value of FRM data prior to CEP coding. In that year, schools with at least 40 percent of students identified as DC had, on average, 79 percent of students coded as FRM-eligible.

3. Data

Our analysis is based on student-level administrative microdata provided by the Missouri Department of Elementary and Secondary Education (DESE). The data panel covers the school years 2011-12 through 2016-17 and thus includes years before and after the implementation of the CEP, which was first adopted by Missouri schools during the 2014-15 school year (hereafter we refer to school years by the spring year; e.g., 2014-15 as 2015). The CEP status for all schools in the state in each year is available from DESE.

We assess the informational content of FRM and DC data using predictive models of student attendance and student achievement in math and English language arts (ELA) in grades 3-8.[1] We define the attendance rate as the total number of days attended divided by the total number of days enrolled, on a 0-1 scale.[2] All test scores are standardized to have a mean of zero and a variance of one within subject-grade-year cells.

Figure 1 documents the rollout of the CEP in Missouri during our data panel. The changes over time in CEP implementation are cumulative and shown as (1) the count of schools, (2) the share of schools (again, with at least one grade in the 3-8 range), and (3) the share of enrollment. The enrollment share is consistently below the share of schools, reflecting the fact that the average CEP-adopting school in Missouri is smaller than the average school statewide. This, in turn, reflects the fact that many eligible schools are in rural areas.

[1] We focus on grades 3-8 due to the statewide testing of students in these grades in math and ELA over the course of our data panel. The attendance models focus on the same grades to ensure that comparisons across the models are not confounded by changes to the composition of the sample.

[2] For students who are enrolled in more than one school in a given year, attendance is calculated across all schools in which they enroll.

Our data include race and gender information for each student, whether the student is an English language learner (ELL), and whether the student has an individualized education program (IEP). We also know each student's FRM eligibility status in each year; moreover, for the years 2013 to 2017, we know each student's DC status. DC data are not yet ubiquitous in state data systems but are increasingly available (Chingos, 2018).

We aggregate FRM and DC data to provide contextual information about the school attended by each student. For example, the share of FRM-eligible students at the school gives information beyond what is conveyed by a student's own FRM-eligibility status (Ehlert et al., 2016). We also construct individual-level panel measures of FRM and DC status that we include in some models. These measures capture the fraction of years—up to and inclusive of the current year—that a student is coded as either FRM-eligible or directly certified in the Missouri data. Michelmore and Dynarksi (2017) show that panel measures provide more information about student disadvantage than contemporaneous measures alone.[3]

Table 1 provides summary statistics for the data, which include over 1,700 schools with at least some coverage of tested grades and subjects (e.g., K-5, K-8, 6-8, etc.) and 1.8 million student-year observations (summed over the pre- and post-CEP years of the data panel). On average during the post-CEP portion of the data panel, 11.6 percent of students in Missouri attended a CEP school.

4. Methodology

4.1 The Effect of CEP-Coding on the Informational Content of FRM Data

We begin with basic models designed to determine how much the CEP has degraded the proxy value of FRM-eligibility as an indicator of student disadvantage. Our initial models do not

[3] We use fractional measures for the panel variables, instead of counts of FRM- or DC-eligible years, to improve comparability of the panel variables across grade levels. For example, the meaning of a count variable for a grade-3 student will not be the same as for a grade-8 student, whereas the fractional measures are more comparable. Like with contemporaneous FRM, the panel FRM variable is influenced by the introduction of the CEP; and the panel DC variable is influenced by our truncated DC data panel. We examine the sensitivity of our findings to these data issues below and find that they do not affect our findings substantively.

consider DC data. Instead, we focus on the "business as usual" setting in modern research and policy applications, which does not incorporate other disadvantage measures. We focus only on contemporaneous FRM information to begin with, which is consistent with how information is typically used by researchers and policymakers. Later on we expand our analysis to include the panel measures of disadvantage.

The initial regressions take the following form:

$$Y_{igst} = \delta_0 + FRM_{it}\psi_1 + \mathbf{X_{it}}\boldsymbol{\psi_2} + \omega_g + \xi_t + e_{igst} \quad (1)$$

$$Y_{igst} = \delta_0 + FRM_{it}\delta_1 + \overline{FRM}_{st}\delta_2 + \mathbf{X_{it}}\boldsymbol{\delta_3} + \overline{\mathbf{X}}_{\mathbf{st}}\boldsymbol{\delta_4} + \lambda_g + \phi_t + \varepsilon_{igst} \quad (2)$$

Equations (1) and (2) are nearly identical; the difference is that equation (2) includes school-average student characteristics as additional predictors of outcomes (to capture educational context). In both equations, Y_{igst} is the outcome of interest—either a math test score, an ELA test score, or the attendance rate (again, on a 0-1 scale)—for student *i* in grade *g* at school *s* in year *t*. FRM_{it} is an indicator equal to one if student *i* is coded as FRM-eligible in year *t*, and in equation (2), $\overline{FRM}_{st}$ is the share of students attending school *s* in year *t* who are coded as FRM-eligible. X_{it} and $\overline{X}_{st}$ are analogous vectors of the other student and school-aggregated characteristics. The student characteristics are as shown in Table 1 and include student race/ethnicity and gender indicators, along with indicators for whether the student is learning English as a second language (ESL) and has an individualized education program (IEP). Conceptually the variables in the *X*-vectors are no different than the FRM-eligibility variables, but we separate out FRM_{it} and $\overline{FRM}_{st}$ visually because the coefficients ψ_1, δ_1, and δ_2 are focal to our analysis. Finally, ω_g / λ_g, ξ_t / ϕ_t and $e_{igst} / \varepsilon_{igst}$ are grade fixed effects, year fixed effects, and idiosyncratic errors clustered at the school level, respectively.[4]

We initially estimate these equations separately using data from the pre- and post-CEP periods for each student outcome. We report on changes to the coefficients ψ_1, δ_1, and δ_2; and changes to the overall predictive power of the models. As noted above, a limitation of the simple pre-post analysis is that other factors may also be changing over the timespan during which the CEP has been adopted in Missouri, which could influence the results.

In order to isolate the effect of CEP we extend the use of the models to the above-described pseudo-coding scenarios. In the first of these we identify all schools that adopted the CEP during the first year in Missouri (2015). Call these "group A" schools. We estimate equation (1) using data from pre-CEP years only (i.e., up through 2014), but code the data as if group-A schools had already adopted the CEP. Noting that no school had actually adopted the CEP during the pre-CEP years, group-A schools are "pseudo-coded" to have adopted the CEP prior to actual adoption.

We estimate equations (1) and (2) using the same exact data, with and without the CEP pseudo-coding, to assess the data-quality consequences of the CEP holding all else equal. Like with the simple pre-post comparison, we focus our attention on the coefficients ψ_1, δ_1, and δ_2, and each model's overall predictive power. Given the basic statistics of measurement error, we expect ψ_1, δ_1, and δ_2 to attenuate toward zero and the overall predictive power of the model to decline with the CEP pseudo-coding, but the magnitudes of these changes are difficult to predict *a priori*.

We build on the above-described scenario with two more-pronounced scenarios. In the first of these, we pseudo-code all schools that ever adopted the CEP at any point during our data panel. Based on the slow growth in CEP adoptions after 2015 illustrated by Figure 1, we do not expect this

[4] Because tests are standardized within subject-grade-year, the grade and year fixed effects are of no practical importance in the models, but we include them for completeness.

change to have a significant impact on the findings. For the final scenario we use school-level DC data to identify all CEP-eligible schools based on the DC student share in 2016 (the middle year of the post-CEP portion of our data panel), then pseudo-code all of these schools as adopters in the pre-CEP period. This last scenario gives the upper-bound effect of the CEP in the sense that it allows for the maximum number of schools to participate as permitted by program rules.[5]

4.2 *The Effect of CEP-Coding on School Accountability*

We also consider the implications of CEP-induced data changes for school accountability policies based on value-added. We estimate school value-added to math and ELA achievement in grades 3-8 using a two-step model following Ehlert et al. (2016) and Parsons, Koedel, and Tan (forthcoming).[6] Specifically, we estimate the following equations sequentially:

$$Y_{igst} = \gamma_0 + Y_{i(t-1)}\gamma_1 + FRM_{it}\gamma_2 + \overline{FRM}_{st}\gamma_3 + \mathbf{X_{it}}\boldsymbol{\gamma}_4 + \overline{\mathbf{X}}_{\mathbf{st}}\boldsymbol{\gamma}_5 + \pi_g + \varsigma_t + \eta_{igst} \quad (3)$$

$$\eta_{igst} = \boldsymbol{\theta}_{\mathbf{s}} + \tau_{igst} \quad (4)$$

The variables in equation (3) overlap entirely with the variables in equation (2) and are defined as above. The only change is that equation (2) models test-score levels (plus attendance, which we do not consider here), whereas equation (3) models test-score growth by including lagged achievement on the right-hand side. There are a variety of ways to control for lagged achievement that are used in

[5] There is selection into CEP adoption conditional on eligibility. However, the results of our analysis indicate that this is a second-order issue because even in the upper-bound pseudo-coding scenario, the effect of the CEP is generally modest and in this scenario endogenous adoptions are irrelevant (given that we code all eligible schools as CEP adopters). With regard to the other scenarios, the selection is relevant. CEP adopters on average have a higher DC share than non-adopters among the pool of eligible schools. For example, using 2014 data, the average DC share at schools that adopted the CEP is 2015 is 0.61, versus 0.50 for eligible schools that did not adopt the CEP. This is consistent with program incentives in that meal costs are reimbursed at a higher rate for schools with a higher fraction of DC students (over the range of DC-share values of 0.400-0.625, per above). If students who attend schools with a higher DC share are more disadvantaged (as indicated by the analysis below), this type of selection will reduce the amount of information degradation of FRM data because the students who experience a status change are disadvantaged relative to the eligible pool of non-FRM students. The pseudo-coding scenarios based on actual CEP adoptions are inclusive of this selection, which is the most informative approach given that this selection is directly incentivized by CEP program rules.

[6] Missouri uses a structurally similar two-step model in its state accountability system, although the state model does not include all of the control variables we use (e.g., see Ehlert et al., 2014).

the literature (Koedel, Mihaly, and Rockoff, 2015); our models use just the same-subject lagged score.[7] The estimates of school value-added are obtained from the vector $\hat{\boldsymbol{\theta}}_{\mathbf{s}}$ taken from equation (4), to which we apply *ex post* empirical Bayes' shrinkage using the procedure described by Koedel, Mihaly, and Rockoff (2015).

The CEP will affect school rankings based on value-added by changing which students are coded as FRM-eligible. This will factor into the regression adjustment in equation (3). We expect CEP coding to positively affect value-added rankings for affected schools. This is because the coefficients γ_2 and γ_3 in equation (3) are negative (despite modest CEP-induced attenuation), and some students who would not be FRM-eligible based on their own circumstances will be coded as FRM-eligible at these schools under the CEP (which affects both individual student coding and the school share). The end result is that the model will predict students who attend CEP schools to score lower than it otherwise would. When they do better than predicted, the difference will be attributed to the school.[8]

4.3 Comparing FRM and DC Data

Next we explore the viability of using DC data either in place of, or in addition to, FRM data in the post-CEP era (i.e., for the years 2015 to 2017). We also expand our use of the disadvantage metrics to include the panel DC and FRM variables. This extension is motivated by Michelmore and Dynarski (2017) and aligns with our goal of broadly evaluating the capacity to identify student disadvantage with data available in the state system.

[7] We have confirmed that our findings are similar if we include lagged off-subject test scores in the VAMs or use polynomials of the lagged scores, as has been done in some recent research (e.g., Chetty, Friedman and Rockoff, 2014).
[8] The issues described here are fundamentally similar if a one-step value-added model is used instead of the two-step model shown in equations (3) and (4). In the case of a one-step model, it would use within-school FRM-eligibility variation to identify the equivalents of γ_2 and γ_3 (for the latter, the only within-school variation is over time). Schools with fixed CEP status will contribute no variation to the identification of these parameters. The model will similarly "over-adjust" predicted achievement for incorrectly-coded FRM students at CEP schools, thereby benefiting these schools.

Our full model for this portion of the analysis is as follows:

$$\begin{aligned}Y_{igst} &= \beta_0 + \mathbf{X_{it}\beta_1} + \mathbf{\overline{X}_{st}\beta_2} + FRM_{it}\beta_3 + \overline{FRM}_{st}\beta_4 + FRM_{it}^{P}\beta_5 + \\ &DC_{it}\beta_6 + \overline{DC}_{st}\beta_7 + DC_{it}^{P}\beta_8 + \rho_g + \tau_t + u_{igst}\end{aligned} \quad (5)$$

Equation (5) is structurally similar to equation (2) but contains more information. Y_{igst}, X_{it}, and $\overline{X}_{st}$ are as defined above. For each disadvantage measure, there are now three variables—for FRM these are FRM_{it}, $\overline{FRM}_{st}$, and FRM_{it}^{P}. The first two variables are as defined in equation (2), and the third is the panel measure. Analogous sets of variables are included for DC status. ρ_g and τ_t in equation (5) denote grade and year fixed effects, and u_{igst} is the idiosyncratic error clustered at the school level.

A complication with the panel FRM variable in the post-CEP era is that it is influenced by the timing of the introduction of the CEP because the CEP affects FRM coverage. Also, the DC panel variable is influenced by the fact that DC data are first available to us in 2013 (we construct the panel DC variable as the share of years between 2013 and *t* in which a student is coded as DC). We examine the sensitivity of our findings to these data issues in the appendix (Appendix Table A.9) by analyzing just the last year of the data panel (2017) when the CEP effect would be most pronounced on the panel FRM variable, and the DC panel variable is most complete. The results are very similar to what we report in the text below using all three post-CEP years. Therefore, we conclude that these data issues do not substantively affect our findings.

5. Results

5.1 The Effect of the CEP

Table 2 shows results from the math-achievement version of equation (1) using pre- and post-CEP data, and the pseudo-coded data. Models with and without the *X*-vector are included for each condition. The results for ELA and attendance are substantively similar to the results in Table 2 with respect to the implications of the CEP—although we can explain much less of the total variance in student outcomes in the attendance models—and thus for ease of presentation we relegate them to the appendix (Appendix Tables A.1 and A.2).

In addition to showing the regression results, Table 2 also shows how the FRM-eligible share of students in Missouri evolves under the various CEP conditions. The first two conditions in the table, for the pre- and post-CEP years of our data panel, show that the percent of CEP-adopting schools grew from 0 in the pre-CEP period to an average of 15.1 percent of schools during the post-CEP period. But this increase in CEP-adopting schools corresponds to a much smaller increase in the share of Missouri students coded as FRM-eligible—just 1.7 percentage points. As noted above, there are two primary reasons for the small increase: (1) CEP-adopting schools typically have a small fraction of non-FRM-eligible students owing to program rules, and (2) the average CEP-adopting school is smaller than the average school in Missouri.

Next we turn to the pseudo-coded conditions. The first pseudo-coded condition also shows an increase in the FRM-eligible student share of 1.7 percentage points, to 52.9 percent. The match with the pre/post comparison is coincidental, likely reflecting a combination of there being more CEP schools on average in the full post-CEP period, offset by improving economic conditions statewide over time from the pre- to post-CEP years (which affects the statewide FRM take-up rate). The second pseudo-coded condition, in which we pseudo-code all schools that ever adopted the CEP by the end of our data panel, only marginally increases the shares of CEP schools and FRM-coded students (to 16.4 and 53.5 percent, respectively), as predicted based on Figure 1.

Finally, in the third condition in columns (9) and (10) we pseudo-code all students in CEP-eligible schools. This gives an upper bound on the effect of the CEP. While just over 30 percent of Missouri schools are CEP eligible, even at this upper-bound the hypothetical effect of the CEP on the share of FRM-coded students in Missouri is modest. The share of FRM-eligible students statewide rises 5.3 percentage points from the pre-CEP baseline, to 56.5 percent. This small increase, even in the upper-bound condition, foreshadows our modest findings with regard to the effect of the CEP on FRM status as a proxy for student disadvantage.

Turning to the regression results, the top row of Table 2 shows estimates of ψ_1 for each condition from models with and without the other control variables. The CEP has essentially no effect on the estimated achievement gap by student FRM status. For example, consider a comparison of the estimates of ψ_1 in columns (1) and (9), which captures the maximum scope for effect of the CEP. These models condition only on the grade and year, and thus the output can be interpreted as what is effectively raw differences in achievement by student FRM status. The results show that without CEP coding in place in column (1), FRM-eligible students score 0.623 standard deviations lower in math than ineligible students on average. With maximum CEP coding in place in column (9) the gap decreases, but only by a negligible 0.014 standard deviations, to 0.609.

When we include all of the other control variables the model is able to better predict student outcomes (i.e., the R-squared increases substantially), but the story with respect to the predictive power of FRM status is essentially unchanged. The analogous comparison of columns (2) and (10) indicates that the CEP reduces the fully-conditioned FRM gap in math achievement by just 0.015 standard deviations; from 0.442 to 0.427. Finally, the changes to the overall predictive power of the models tell a similar story. Sticking with the comparison of columns (2) and (10), the reported R-squared values show that the CEP, at its upper bound effect, reduces the variance in math achievement explained by the model by just 0.4 percentage points (from 22.9 to 22.5 percent).

Table 3 follows the structure of Table 2 but shows output from equation (2) with the school-aggregated variables included. Again, we show results for math achievement in the main text and relegate the findings for ELA achievement and attendance to the appendix because of their similarity (Appendix Tables A.3 and A.4). Like in Table 2, in Table 3 there is no discernable change in the test-score gap between individual students who differ by FRM coding status across the pseudo-coded scenarios, conditional on the school FRM share. In fact, $\hat{\delta}_1$ becomes nominally *more negative* as CEP coverage increases. In isolation this result is directionally inconsistent with the CEP inducing attenuation bias in $\hat{\delta}_1$. However, it is clear that there is an attenuating effect of the CEP loading onto $\hat{\delta}_2$. In both the sparse models (i.e., the models without the *X*-vectors) and the full models, $\hat{\delta}_2$ consistently declines as the influence of the CEP increases. Comparing columns (1) and (2) to columns (9) and (10) reveals a sharp change—the magnitude of $\hat{\delta}_2$ is reduced by roughly half. For example, the sparse model in column (1), using the real data, indicates that moving from a school with 0 to 100 percent FRM eligibility is associated with a reduction in test-score performance of 0.763 student standard deviations, whereas the same change in column (9) is associated with a reduction in student performance of just 0.391 standard deviations.

The effect of the CEP in Table 3—driven by the effect on the school-aggregated FRM variable—can be illustrated by comparing the predicted test score gap between two hypothetical students: (1) a student who is not coded as FRM-eligible attending a school where 25 percent of students are FRM-eligible, and (2) a student who is FRM-eligible attending a school where 75 percent of students are FRM-eligible. Per Table 1, note that the 50 percent gap in the share of FRM eligible students in a school corresponds to roughly two standard deviations in the distribution of that variable. Based on the results from the model in column (1), the estimated achievement gap between these students is 0.849 standard deviations of student achievement (0.467 + 0.5*0.763). But based on the model in column (9), the estimated gap is just 0.670 standard deviations (0.475+0.5*0.391). These results make clear that the CEP reduces the level of disadvantage conveyed by the share of FRM-eligible students at a school.

The evolution of the R-squared values in Table 3 is similar to what we find in Table 2, reinforcing the finding that the CEP has a modest effect on the ability of FRM data to explain variation in mathematics test scores. Focusing on the comparison between the sparse models in columns (1) and (9), the total change in the R-squared induced by the CEP is 0.021. The change is about half as large when we use the full model in columns (2) and (10) (0.012) because the X vectors partially compensate for the information loss in FRM data due to the CEP.

Selected results highlighting the key findings from Tables 2 and 3 are presented visually in Figure 2.

Above we suggest two mechanisms to explain the generally modest effect of the CEP on the informational content of FRM data. The first is that the CEP changes FRM status for students who attend high-poverty schools, which limits the substantive impact of technical inaccuracies (i.e., these students are still disadvantaged even if they are not FRM eligible individually). The second is that relatively few students experience a change in status as a result of the CEP, as shown in Tables 2 and 3. To disentangle these competing mechanisms, we take the same total change in FRM status induced by the CEP in the upper-bound scenario—5.3 percent of Missouri students in pseudo-coding scenario 3—but instead of pseudo-coding schools that are eligible for the CEP based on their DC shares, we randomly select the schools and code them as if they had implemented the CEP. In this counterfactual, the scope of the CEP is held constant in terms of the number of students who experience a change in FRM status, but the students who experience a change are no longer concentrated in high-poverty schools.

The analysis reveals that both mechanisms play a role in limiting the impact of the CEP, but the more important factor is the small number of students who experience a status change. When 5.3 percent of Missouri students are switched from FRM=0 to FRM=1 by randomly switching the CEP-coding status of schools, the estimated achievement gaps by FRM status and the school FRM share remain fairly close to what we report in Tables 2 and 3. The results from this exercise are provided in Appendix Table A.5.

5.2 School Accountability

Table 4 shows results from value-added models of math achievement for students in grades 3-8 (substantively similar results for ELA are available in Appendix Table A.6). To illustrate the effect of the CEP, we report the average percentile ranking of schools that actually adopted the CEP in the first year in Missouri under the different conditions. We also report the number of these schools that are in the top quintile of value-added rankings. Results are not shown for the other pseudo-coded conditions because inference from Table 4 is confounded if the number of focal schools changes across columns. The table is structured so that in each column, ranking outcomes for the same set of schools are reported.

Although the previous section suggests a generally modest effect of the CEP on the total explanatory power of FRM data, Table 4 shows that the CEP boosts estimated value-added for participating schools. As discussed in Section 4, the mechanism is that for a school that moves to CEP status, the predicted performance of re-coded FRM students declines per equation (3). This is because more of the students at these schools are coded as FRM eligible and the FRM-eligible school share goes to 1.0 (the coefficients on both of these variables in equation (3) are negative). Correspondingly, the actual performance of these students relative to predicted performance improves. This benefits the school in the value-added calculation.

Comparing the results in the first and last columns of Table 4, where inference is cleanest, the average percentile rank of year-1 CEP adopters in the actual pre-CEP data is 48.8. But with the

CEP coding rules, the average percentile rank rises just over six points, to 54.9. Similarly, the number of CEP adopters in the top quintile rises from 48 to 65 because of the CEP coding.

Whether the CEP-induced shift in rankings is desirable is an open question. On the one hand, at a fundamental level it is driven by a data inaccuracy, which makes it unappealing. However, it could be viewed positively because in states with accountability systems that incorporate value-added, it gives a clear incentive for schools to adopt the CEP. This should be appealing to state education agencies for several reasons. First, the CEP is a federally funded program so states interested in expanding access to federal aid for their schools will be supportive of increased take-up of the CEP. Second, the literature on universal free meals, although nascent, suggests that students benefit from these programs academically and otherwise (Dotter, 2013; Gordon and Ruffini, 2018; Schwartz and Rothbart, 2017). Finally, Parsons, Koedel, and Tan (forthcoming) show that even value-added models that take great care to avoid bias favoring advantaged schools—like the two-step model described by equations (3) and (4)—remain at least marginally biased in favor of these schools under the most common estimation conditions. The CEP's positive effect on the rankings of low-income schools could offset some of this bias.

5.3 Comparing FRM and DC Data

Next we shift our focus to compare the predictive power of FRM and DC data in the post-CEP era (i.e., from 2015 to 2017). Table 5 shows results from versions of equation (5) where math achievement is the dependent variable and combinations of contemporaneous FRM and DC controls are included—we do not include the panel variables initially. The first two columns show models that include individual student FRM and DC variables separately. These models allow for a clean comparison of the achievement gaps predicted by these measures without any other controls. The results show that even in the post-CEP era, students coded as FRM-eligible have lower test scores on average than DC-coded students. Specifically, the test-score gap between FRM and non-

FRM students is 0.651 student standard deviations, whereas the analogous gap between DC and non-DC students is 0.590.

Columns (3)-(4) add the school aggregates. Some of the weight on the individual measures shifts to the aggregate measures and the overall explanatory power of both models improves. The model using FRM data remains modestly more predictive of student achievement. One notable result is that the school-average DC share is a much stronger predictor of test scores than the school-average FRM share. This is driven in large part by the fact that there is less variation in the DC school share per Table 1. Specifically, the student-weighted standard deviation of the FRM school share from 2015-2017 is 0.256, whereas for the DC share it is 0.171. The implication is that the effect of a move from 0-100 percent coverage, which is what the coefficients capture, represents a lager change in the distribution for the DC share. This is much less of an issue with the individual FRM and DC controls, which have similar variances (again, see Table 1). While differential variance in the school-aggregated variables is a partial explanation, the gap in the coefficients is more than would be expected if this was only a matter of distributional rescaling. The CEP-induced reduction in the informational content of school-aggregated FRM data, as documented in the previous section, is also surely a factor.

Finally, column (5) shows results from a model that includes both the FRM and DC variables together (individual and aggregate), and column (6) further adds the other control variables in the individual and school-aggregated X vectors. The loading on the coefficients of interest in columns (5) and (6) follows from the previous columns of the table. That is, individual FRM is more predictive of student achievement than individual DC, but school-average DC is more predictive than school-average FRM. A summary takeaway from column (5) is that the total explanatory power of the model when both types of measures of disadvantage are included (R-squared: 0.143) is not substantially above the total explanatory power when either FRM (R-squared: 0.128) or DC (R-squared: 0.123) information is included in isolation. There is additional explanatory power to be gained by combining both types of measures, but the gain is modest.

As above, we also replicate the analysis in Table 5 for ELA achievement and obtain similar results, which are provided in Appendix Table A.7. However, unlike with the preceding analyses, when comparing the FRM and DC data there are some differences in the results when attendance is the outcome of interest. Therefore, we show results for the attendance models in Table 6. First, Table 6 shows that FRM and DC data are less predictive of attendance than they are of student achievement. This is readily apparent from the R-squared values reported at the bottom of the table.[9] With respect to our investigation of the relative predictive power of FRM and DC data, Table 6 shows that whereas FRM data are marginally more predictive of achievement than DC data, the reverse is true for attendance.

Next, in Tables 7 and 8 we add the panel measures of disadvantage to the models. We continue to relegate the ELA results to the appendix (Appendix Table A.8) because of their similarity to the math results. We also suppress the coefficients for the contemporaneous individual and school-aggregate FRM and DC measures for presentational convenience. A notable feature of the models that is not observable given this suppression is that the panel variables, in addition to improving the predictive power of the models overall, predominantly take the weight off the individual-contemporaneous FRM and DC variables. This is intuitive because with the panel variables in place, the individual-contemporaneous measures capture only the marginal value of current status conditional on cumulative status. It is also consistent with the evidence presented by Michelmore and Dynarski (2017) emphasizing the relative importance of cumulative measures of disadvantage.

[9] This is also the case in the earlier analyses and can be seen by comparing Tables 2 and 3 to their attendance analogs in the appendix.

While it is true that the panel variables increase the predictive power of the models, the increase is generally modest, especially in the richest specifications. For example, the full model in Table 5 explains 24.4 percent of the variance in math achievement, whereas the full model in Table 7—inclusive of the panel variables for each disadvantage measure—explains 25.3 percent of the variance, less than one percentage point more. A substantively similar pattern is presented in Tables 6 and 8 for the attendance outcome. While the limited gains in explanatory power afforded by the panel variables is somewhat surprising, an explanation is that using multiple measures of disadvantage serves largely the same function as using panel variables of the same measure—i.e., to provide a more complete picture of student disadvantage.[10]

We conclude with an accounting of implied achievement and attendance gaps between students who differ in various ways by FRM and DC conditions based on the results in Tables 5-8. The most basic comparison—between students individually coded as either FRM or DC eligible—per the first two columns in Tables 5-6, is straightforward. We also extend our analysis to compare students who differ by measured gaps in the school aggregates and panel variables. Because the DC and FRM aggregate and panel variables differ in their distributions, we consider two types of comparisons: one based on absolute changes in these variables and another based on distributional changes. Specifically, we compare FRM- and DC-based gaps that differ by the individual student's own coded status, plus either a 0.50 change in the school-average share (i.e., representing a move between schools that are 25 and 75 percent FRM or DC) or a one-standard deviation change in the school-average share (i.e., 0.256 for FRM and 0.171 for DC, per Table 1). We then further compare

[10] Recall that the informational content of the panel variables changes over time as the CEP persists, and as we get further away from the first year of DC data (2013). We examine the sensitivity of our findings to these data issues in Appendix Table A.9 (focused on the math achievement models) by analyzing just the last year of our data panel when the CEP effect would be most pronounced, and the DC data most complete (2017). The results are very similar to what we report in the text using the entire post-CEP portion of the data panel.

students who differ by 1.0 in the FRM or DC panel variables and by one standard deviation of these variables, respectively (in this case the standard deviations are much closer per Table 1: 0.443 for FRM and 0.430 for DC).[11]

In summary, we make the following five comparisons between:

1. Students who differ by individual FRM or DC coded status (based on estimates from columns (1) and (2) of Tables 5 and 6).
2. Students who differ by individual FRM or DC coded status, and by *0.50* in the share of students at the school coded as either FRM or DC (based on estimates from columns (3) and (4) of Tables 5 and 6).
3. Students who differ by individual FRM or DC coded status, and by *one standard deviation* in the share of students at the school coded as either FRM or DC (based on estimates from columns (3) and (4) of Tables 5 and 6).
4. Students who differ by individual FRM or DC coded status, by *0.50* in the share of students at the school coded as either FRM or DC, and by *1.0* in the panel FRM or DC measure (based on estimates from columns (1) and (2) of Tables 7 and 8).
5. Students who differ by individual FRM or DC coded status, by *one standard deviation* in the share of students at the school coded as either FRM or DC, and by *one standard deviation* in the panel FRM or DC measure (based on estimates from columns (1) and (2) of Tables 7 and 8).

Table 1 makes clear that any fixed-distance move in the DC and FRM distributions will reflect a larger move in the DC distribution. Comparatively, it may be desirable to use DC data to identify a more targeted group of disadvantaged students. However, FRM data can be used to cast a wider net. The comparisons in the five scenarios described above cover the most important possibilities.

Results for the comparisons are shown in Figures 3 and 4 for mathematics achievement and attendance, respectively; again based on estimates from Tables 5-8.[12] We also illustrate changes to the explanatory power of the models as we add more variables within each type of measure (FRM or DC). For math achievement, Figure 3 visualizes the broad point that FRM and DC data are similarly informative about student disadvantage, and if anything FRM data seem generally more informative

[11] For the panel variables, an absolute change of 1.0 is most informative because, as noted previously, most of the weight on the contemporaneous individual FRM and DC variables shifts to the panel variables when they are both included in the specifications simultaneously (results available from the authors upon request).

[12] The comparisons in the figure can be reproduced based on the results shown in Tables 5-8, with the exception that several coefficients in Tables 7 and 8 are suppressed for ease of presentation. These coefficients are available from the authors upon request.

about achievement gaps. Comparison numbers (2) and (4) from above suggest that the achievement gaps by DC are larger, but as is made clear in the analogous comparison numbers (3) and (5), this is due to the aforementioned differences in the standard deviations of the school-aggregate FRM and DC variables.

Figure 4 presents parallel results for attendance rates, over which our models have much less predictive power in total (as noted previously). It is readily apparent that DC data are more effective in predicting gaps in student attendance. However, in terms of policy significance, the differences in the predictive power of FRM and DC data are modest—e.g., note the compressed scale on the vertical axes of the graphs.

6. Discussion and Conclusion

Setting aside the substantive implications of the CEP, there has been much consternation over how it affects the use of FRM data to identify student disadvantage in education research and policy applications. To the best of our knowledge, we present the first comprehensive analysis designed to explore this issue, at least insofar as FRM eligibility relates to consequential student outcomes. On the whole, we interpret our findings as showing that the effect of the CEP on the informational quality of FRM data is quite modest, even when we consider the potential upper bound effect, although some nuance is in order. Specifically, while the CEP has essentially no effect on the level of disadvantage conveyed by individual FRM-eligibility, it does degrade the quality of information conveyed by the FRM-eligible share in a school.

We also perform a comparative analysis of FRM and DC data to determine their relative efficacy in proxying for student disadvantage in the post-CEP era. DC data have been advocated as a substitute for FRM data in several articles and reports that raise concerns over the potential data-quality consequences of the CEP (Camera, 2019; Chingos, 2018; Greenberg, 2018). Our comparative analysis shows that FRM and DC data are similarly informative about student disadvantage in the post-CEP era. We also show that surprisingly little is gained by combining both types of information to improve the identification of disadvantaged students, as evidenced by the very small increase in the explanatory power of our models when we include both types of measures at once, as opposed to either type of measure individually.

For researchers interested in using FRM data in its traditional role as a control to account for individual student disadvantage, our results suggest that the CEP is of limited concern. Because (a) the students whose FRM status is changed by the CEP are already attending high poverty schools (which limits the substantive importance of CEP-induced data inaccuracies) and (b) the CEP does not result in status changes for a large number of students, the individual FRM control performs no worse with the CEP in place than without it. However, for researchers who are also interested in controlling for schooling context using the FRM-eligible school share, our results indicate the quality of this control is degraded by the CEP. To offset some of the information loss researchers can look to buttress their models with other information. A simple suggestion is to add an indicator variable to the model for whether the school adopted the CEP. This will help to offset the effect of over-representation of FRM-eligible students in CEP schools and force identification of the coefficient on school-average FRM to rely on variation provided only from non-CEP schools. Researchers could also pull in additional, related data from non-education sources, such as local area information about household incomes and education levels (e.g., from the U.S. Census), although at the potential cost of coverage gaps or misalignment between outside data and district and school boundary lines.

The implications for policymakers are substantively similar. If the goal is to monitor achievement gaps by individual FRM status statewide, our results suggest that the introduction of the CEP is effectively ignorable for the reasons outlined above. That said, we recognize that FRM data may be conceptually less appealing with the CEP in place; if the continued use of FRM data is unpalatable in some locales, DC data are a viable alternative when available. States that use school-level aggregates to classify schools in their accountability systems and those with ambitions to create more informative measures of student disadvantage, particularly measures that account for local context (e.g., by using school aggregates), will face the same concerns as researchers given the effect of the CEP on the informational content of school-aggregated FRM data.

Finally, we conclude with a brief note about the generalizability of our findings to other states. Perhaps the first-order issues pertaining to generalizability are CEP eligibility and take-up rates. In states where eligibility and take-up rates are similar to eligibility and take-up rates in Missouri, it seems likely that our substantive findings will generalize given the structure of the CEP program. Thus, other states can quickly assess the likely applicability of our findings to their circumstances by producing these basic summary statistics. Our findings may not generalize to states where many more schools are eligible for the CEP, and/or where take-up rates are higher, which would give the potential for a larger scope of the CEP effect. In such cases, our work provides an analytic plan that researchers can follow to assess the implications of the CEP given their own local conditions.

References

Bass, D. N. (2010). Fraud in the lunchroom? *Education Next, 10*(1), 67-71.

Camera, L. (2019). Miscounting Poor Students. *U.S. News & World Report.* Retrieved 01.09.2019 at: https://www.usnews.com/news/education-news/articles/2019-01-07/why-its-getting-harder-to-count-poor-children-in-the-nations-schools

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review* 104(9), 2593-2632.

Chingos, M.M. (2018). A promising alternative to subsidized lunch receipt as a measure of student poverty. Policy report. Washington DC: Brookings Institute.

Chingos, M.M. (2016). No more free lunch for education policymakers and researchers. *Evidence Speaks Reports* 1(20), 1-4. Washington DC: Brookings Institute.

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

Dotter, D. (2013). Breakfast at the Desk: The impact of universal breakfast programs on academic performance. Unpublished manuscript.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2014). The Missouri Compromise. Policy Blog. *Education Next.* Retrieved 12.06.2018 at: https://www.educationnext.org/the-missouri-compromise/

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2016). Selecting growth measures for use in school evaluation systems: Should proportionality matter? *Educational Policy* 30(3): 465-500.

Gordon, N.E. & Ruffini, K.J. (2018). School nutrition and student discipline: Effects of schoolwide free meals. National Bureau of Economic Research Working Paper No. 24986.

Greenberg, E. (2018). New measures of student poverty: Replacing free and reduced-price lunch status based on household forms with direct certification. Education Policy Program policy brief. Washington DC: Urban Institute.

Harwell, M., & LeBeau, B. (2010). Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher* 39(2), 120-131.

Koedel, C., Mihaly, K., & Rockoff, J. (2015). Value-added modeling: A review. *Economics of Education Review* 47, 180-195.

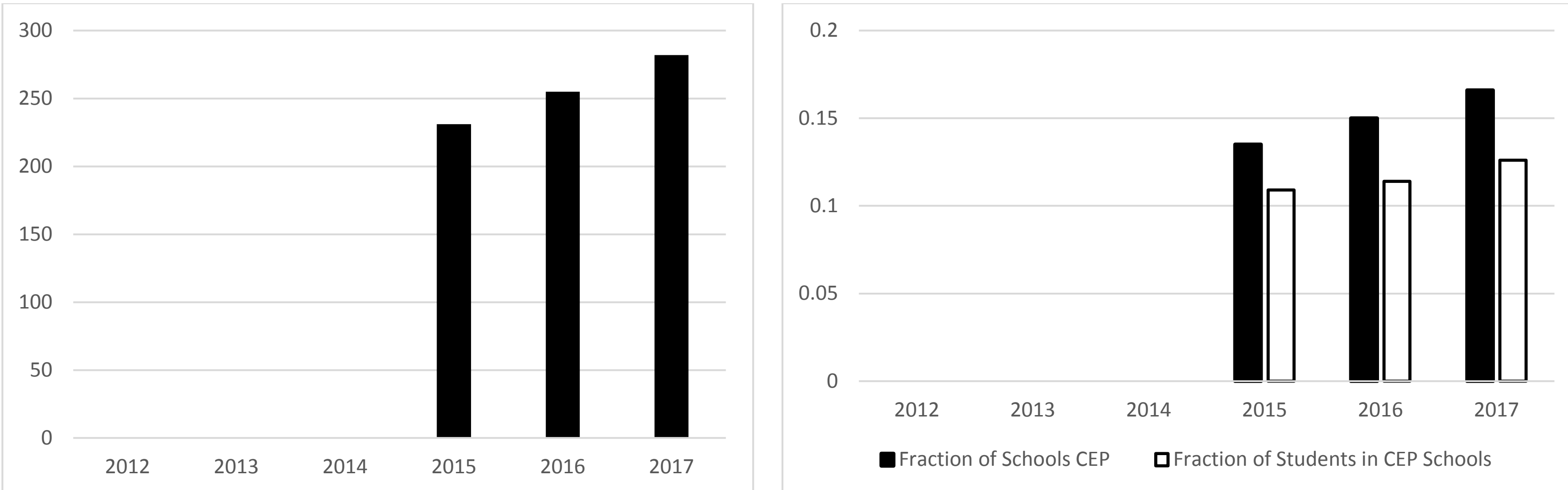
Michelmore, K., & Dynarski, S. (2017). The gap within the gap: Using longitudinal data to understand income differences in educational outcomes. *AERA Open* 3(1), 1-18.

Parsons, E., Koedel, C., & Tan, L. (forthcoming). Accounting for student disadvantage in value-added models. *Journal of Educational and Behavioral Statistics.*

Riddle, W. (2015). Implications of Community Eligibility for the Education of Disadvantaged Students Under Title-I. Policy Report. Washington DC: Center on Budget and Policy Priorities.

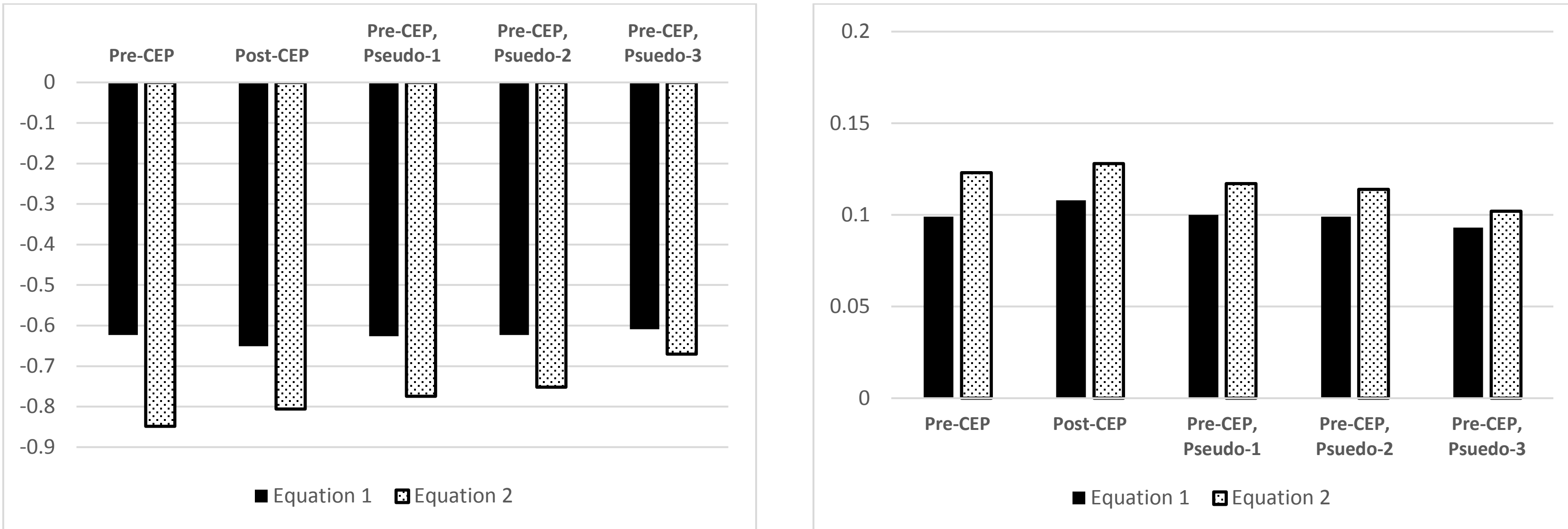
Schwartz, A.E., & Rothbart, M.W. (2017). Let them Eat Lunch: The Impact of Universal Free Meals on Student Performance. Maxwell School Center for Policy Research Working Paper #203.

Figure 1. CEP School Counts, and CEP Coverage of Schools and Students, in Missouri Over Time for Schools with any Combination of Grades 3-8.



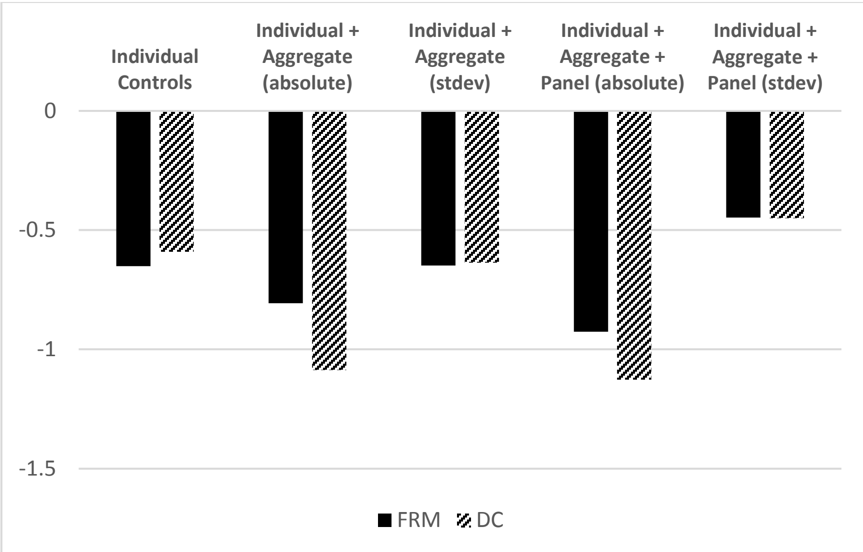
Notes: The graph on the left shows the number of schools with any combination of grades 3-8 (thus in our analytic sample) implementing the CEP in each year. The graph on the right shows CEP schools as a fraction of all eligible schools (with the same gradespan restriction), and the corresponding fraction of students in covered schools. All representations are cumulative—i.e., the numbers in 2016 reflect the cumulative effect of adoptions in 2015 and 2016. As in the main text, school years are indicated by the spring year.

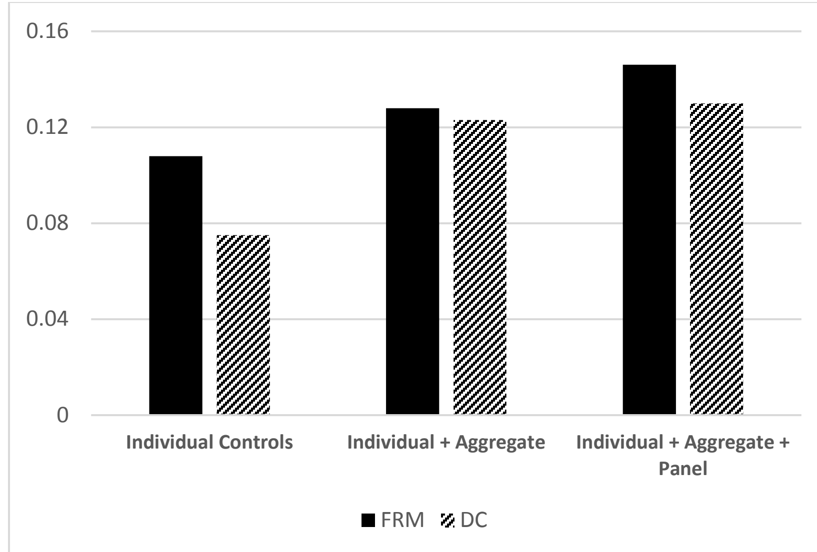
Figure 2. Predicted gaps in math achievement between students who differ by FRM status and FRM school conditions (left), and the overall predictive power of the sparse math achievement models shown by equations (1) and (2) (right), under various CEP conditions.



Notes: The graph on the left shows math achievement gaps as estimated by sparse versions of equations (1) and (2) (i.e., without the X-vector controls) under various CEP conditions. The gaps from equation (1) compare a FRM-eligible student to an ineligible student. The gaps from equation (2) compare students who differ by own FRM eligibility and have a 50 percentage point gap in the FRM eligibility shares at their schools (i.e., $\Delta\overline{FRM} = 0.50$, which is roughly a two-standard-deviation change in the distribution of the school FRM-eligible share). The graph on the right shows R-squared values from the sparse versions of equations (1) and (2), which indicate the overall predictive power of the models over math achievement under the various CEP conditions.

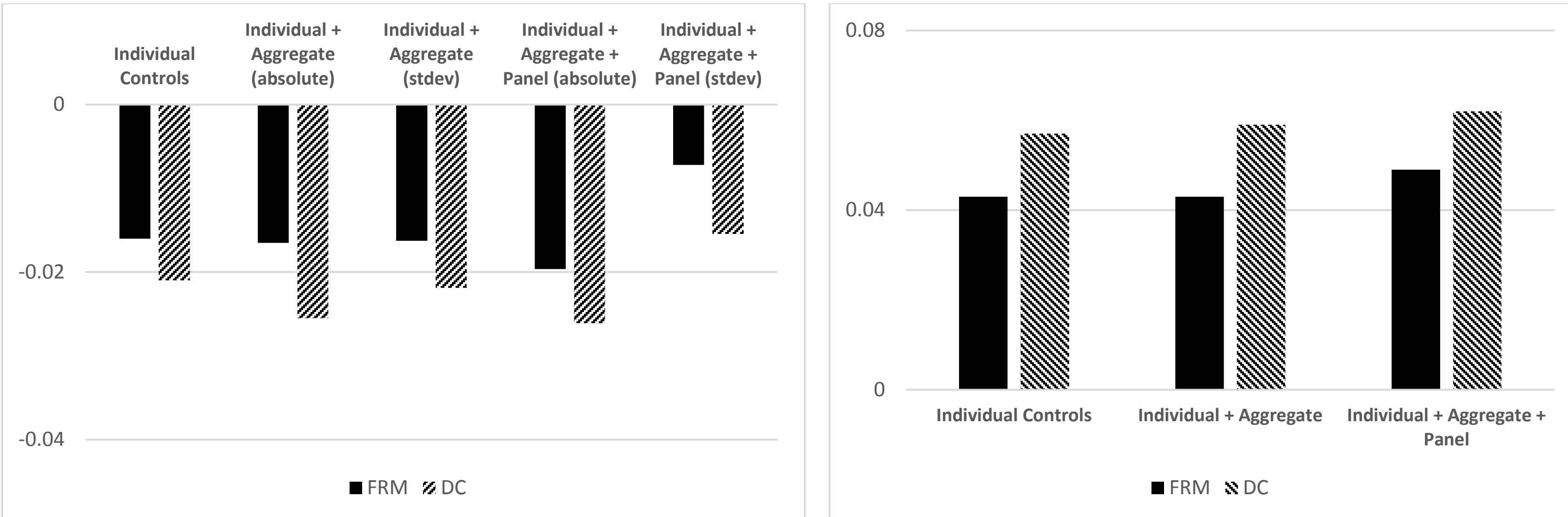
Figure 3. Predicted gaps in math achievement between students who differ by various measures of FRM and DC (left), and the overall predictive power of versions of the math achievement model shown in equation (5) using FRM versus DC data (right), during the post-CEP era.





Notes: The graph on the left shows math achievement gaps estimated using FRM or DC information from versions of equation (5). We make five comparisons as described by Section 5.3 in the text. These are between students who: (1) differ by individual FRM or DC coded status, (2) differ by the individual FRM or DC coded status, and by *0.50* in the share of students at the school coded as either FRM or DC, (3) differ by the individual FRM or DC coded status, and by *one standard deviation* in the share of students at the school coded as either FRM or DC, (4) differ by the individual FRM or DC coded status, by *0.50* in the share of students at the school coded as either FRM or DC, and by *1.0* in the panel FRM or DC measure, and (5) differ by the individual FRM or DC coded status, by *one standard deviation* in the share of students at the school coded as either FRM or DC, and by *one standard deviation* in the panel FRM or DC measure. The graph on the right shows R-squared values from versions of equation (5) that include grade and year fixed effects and either the FRM or DC information indicated by the labels on the horizontal axis.

Figure 4. Predicted gaps in attendance rates between students who differ by various measures of FRM and DC (left), and the overall predictive power of versions of the attendance model shown in equation (5) using FRM versus DC information (right), during the post-CEP era.



Notes: These graphs are analogs to the graphs in Figure 3 but based on the models of attendance rates instead of math achievement. See notes to Figure 3.

Table 1. Descriptive Statistics

	Pre-CEP Years 2012-14	Post-CEP Years 2015-17
Student Outcomes	Mean (stdev)	Mean (stdev)
Standardized Math Score	0.016 (0.989)	0.010 (0.991)
Standardized Reading Score	-0.006 (0.986)	-0.018 (0.988)
Attendance Rate	0.954 (0.046)	0.954 (0.044)
Student Characteristics		
Race/Ethnicity: White	0.743 (0.437)	0.726 (0.446)
Race/Ethnicity: Black	0.164 (0.370)	0.159 (0.366)
Race/Ethnicity: Hispanic	0.050 (0.218)	0.059 (0.236)
Race/Ethnicity: American Indian	0.004 (0.065)	0.004 (0.063)
Race/Ethnicity: Asian/Pacific Islander	0.020 (0.139)	0.020 (0.142)
Race/Ethnicity: Other	0.019 (0.137)	0.031 (0.173)
Female	0.488 (0.500)	0.488 (0.500)
English as Second Language (ESL)	0.030 (0.172)	0.040 (0.196)
Individual Education Program (IEP)	0.123 (0.328)	0.130 (0.336)
Measures of Disadvantage & CEP		
FRM Status (student level)	0.512 (0.500)	0.529 (0.499)
FRM School Share (student weighted)	0.507 (0.227)	0.523 (0.256)
Attends CEP School	0	0.116 (0.321)
Direct Certification Status*	0.279 (0.449)	0.300 (0.458)
Direct Certification School Share* (student weighted)	0.275 (0.165)	0.295 (0.171)
Panel FRM	-	0.528 (0.443)
Panel DC	-	0.319 (0.430)
N (Schools)	1748	1737
N (Student-Years)	920541	916760

Notes: Data on direct certification status are only available—and thus only reported in the table—for the school years 2012-13 to 2016-17. For the analysis, we use the panel variables in the post-CEP years only and thus report descriptive statistics for these variables in just these years.

Table 2. Estimates of the Math Achievement Gap in Grades 3-8 by FRM Coding Status, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.623	-0.442	-0.651	-0.469	-0.626	-0.441	-0.623	-0.438	-0.609	-0.427
	(0.011)***	(0.007)***	(0.013)***	(0.009)***	(0.012)***	(0.008)***	(0.012)***	(0.008)***	(0.013)***	(0.009)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.099	0.229	0.108	0.225	0.100	0.228	0.099	0.228	0.093	0.225
N(Students)	916461	916461	909974	909974	916461	916461	916461	916461	916461	916461

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses.
*** indicates statistical significance at the 1 percent level.

Table 3. Estimates of the Math Achievement Gap in Grades 3-8 by FRM Coding Status of Individual Students and Schools, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.467	-0.346	-0.484	-0.358	-0.474	-0.351	-0.474	-0.351	-0.475	-0.349
	(0.006)***	(0.004)***	(0.006)***	(0.004)***	(0.006)***	(0.004)***	(0.006)***	(0.004)***	(0.007)***	(0.004)***
FRM School Share	-0.763	-0.603	-0.644	-0.539	-0.601	-0.449	-0.556	-0.402	-0.391	-0.255
	(0.036)***	(0.033)***	(0.034)***	(0.038)***	(0.036)***	(0.034)***	(0.036)***	(0.034)***	(0.032)***	(0.028)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.123	0.242	0.128	0.236	0.117	0.237	0.114	0.235	0.102	0.230
N(Students)	916461	916461	909974	909974	916461	916461	916461	916461	916461	916461

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses.
*** indicates statistical significance at the 1 percent level.

Table 4. The Effect of CEP Coding on School Accountability Based on Value-Added to Test Scores in Mathematics, Grades 3-8.

	Pre-CEP	Post-CEP	Pre-CEP with Pseudo-Coding 1
Average percentile ranking in the school distribution of year-1 CEP adopters (in 2014-2015) using VAM for all students	48.8	51.3	54.9
Number of year-1 CEP adopters (in 2014-2015) ranked in the top quintile using VAM for all students	48	56	65

Notes: Ranking outcomes reported for 231 schools that adopted the CEP in 2015 (per Figure 1).

Table 5. Estimated Math Achievement Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17).

	(1)	(2)	(3)	(4)	(5)	(6)
Individual FRM	-0.651		-0.484		-0.380	-0.273
	(0.013)***		(0.006)***		(0.006)***	(0.004)***
Individual DC		-0.590		-0.402	-0.197	-0.164
		(0.011)***		(0.006)***	(0.004)***	(0.003)***
FRM School Share			-0.644		0.096	-0.052
			(0.034)***		(0.068)	(0.075)
DC School Share				-1.369	-1.189	-0.892
				(0.041)***	(0.097)***	(0.114)***
Other Controls						Y
R-Squared	0.108	0.075	0.128	0.123	0.143	0.244
N (Students)	909974	909974	909974	909974	909974	909974

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses. *** indicates statistical significance at the 1 percent level.

Table 6. Estimated Attendance Rate Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17).

	(1)	(2)	(3)	(4)	(5)	(6)
Individual FRM	-0.016		-0.016		-0.007	-0.008
	(0.000)***		(0.000)***		(0.000)***	(0.000)***
Individual DC		-0.021		-0.020	-0.016	-0.015
		(0.000)***		(0.000)***	(0.000)***	(0.000)***
FRM School Share			-0.001		0.0117	0.0170
			(0.001)		(0.002)***	(0.002)***
DC School Share				-0.011	-0.028	-0.019
				(0.002)***	(0.004)***	(0.004)***
Other Controls						Y
R-Squared	0.043	0.057	0.043	0.059	0.063	0.072
N (Students)	916760	916760	916760	916760	916760	916760

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses. *** indicates statistical significance at the 1 percent level.

Table 7. Estimated Math Achievement Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17), Inclusive of Panel Measures.

	(1)	(2)	(3)	(4)
Panel FRM	-0.616		-0.511	-0.415
	(0.013)***		(0.011)***	(0.009)***
Panel DC		-0.445	-0.108	-0.092
		(0.009)***	(0.007)***	(0.006)***
Contemporary Individual and School-Aggregated FRM	Y		Y	Y
Contemporary Individual and School-Aggregated DC		Y	Y	Y
Other Controls				Y
R-squared	0.146	0.130	0.156	0.253
N (students)	909974	909974	909974	909974

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses. *** indicates statistical significance at the 1 percent level.

Table 8. Estimated Attendance Rate Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17), Inclusive of Panel Measures.

	(1)	(2)	(3)	(4)
Panel FRM	-0.017		-0.010	-0.011
	(0.000)***		(0.000)***	(0.000)***
Panel DC		-0.014	-0.008	-0.008
		(0.000)***	(0.000)***	(0.000)***
Contemporary Individual and School-Aggregated FRM	Y		Y	Y
Contemporary Individual and School-Aggregated DC		Y	Y	Y
Other Controls				Y
R-squared	0.049	0.062	0.067	0.077
N (students)	916760	916760	916760	916760

Notes: All models include grade and year fixed effects. Standard errors clustered by school reported in parentheses. *** indicates statistical significance at the 1 percent level.

Appendix Tables

Appendix Table A.1. Estimates of the English Language Arts Achievement Gap in Grades 3-8 by FRM Coding Status, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.607	-0.433	-0.656	-0.482	-0.607	-0.429	-0.604	-0.426	-0.588	-0.412
	(0.010)***	(0.007)***	(0.012)***	(0.009)***	(0.011)***	(0.008)***	(0.011)***	(0.008)***	(0.012)***	(0.008)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.097	0.260	0.115	0.252	0.097	0.258	0.096	0.258	0.090	0.255
N(Students)	918594	918594	914834	914834	918594	918594	918594	918594	918594	918594

Notes: This table replicates the analysis in Table 2 from the main text but using English language arts achievement as the outcome. The notes to Table 2 apply.

Appendix Table A.2. Estimates of the Attendance Rate Gap in Grades 3-8 by FRM Coding Status, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.018	-0.017	-0.016	-0.015	-0.018	-0.017	-0.017	-0.016	-0.017	-0.016
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.050	0.058	0.043	0.051	0.049	0.056	0.049	0.056	0.045	0.053
N(Students)	920541	920541	916760	916760	920541	920541	920541	920541	920541	920541

Notes: This table replicates the analysis in Table 2 from the main text but using the attendance rate as the outcome. The notes to Table 2 apply.

Appendix Table A.3. Estimates of the English Language Arts Achievement Gap in Grades 3-8 by FRM Coding Status of Individual Students and Schools, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.462	-0.340	-0.508	-0.381	-0.467	-0.342	-0.467	-0.342	-0.466	-0.338
	(0.005)***	(0.003)***	(0.006)***	(0.004)***	(0.006)***	(0.004)***	(0.006)***	(0.004)***	(0.006)***	(0.004)***
FRM School Share	-0.709	-0.586	-0.569	-0.505	-0.552	-0.436	-0.510	-0.390	-0.353	-0.247
	(0.030)***	(0.027)***	(0.031)***	(0.035)***	(0.031)***	(0.030)***	(0.031)***	(0.030)***	(0.028)***	(0.024)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.118	0.272	0.131	0.262	0.112	0.266	0.109	0.265	0.097	0.259
N(Students)	918594	918594	914834	914834	918594	918594	918594	918594	918594	918594

Notes: This table replicates the analysis in Table 3 from the main text but using English language arts achievement as the outcome. The notes to Table 3 apply.

Appendix Table A.4. Estimates of the Attendance Rate Gap in Grades 3-8 by FRM Coding Status of Individual Students and Schools, Various CEP Conditions.

	Pre-CEP		Post-CEP		Pre-CEP Pseudo-Coding 1		Pre-CEP Pseudo-Coding 2		Pre-CEP Pseudo-Coding 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FRM	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.017
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
FRM School Share	-0.010	-0.001	-0.001	0.006	-0.006	0.003	-0.005	0.004	-0.001	0.006
	(0.002)***	(0.001)	(0.001)	(0.001)***	(0.002)***	(0.001)***	(0.002)***	(0.001)***	(0.001)	(0.001)***
Other Controls		Y		Y		Y		Y		Y
Share of Students FRM	51.2%		52.9%		52.9%		53.5%		56.5%	
Share of Schools CEP	0		15.1%		13.2%		16.4%		30.7%	
R-Squared	0.052	0.064	0.043	0.055	0.050	0.062	0.050	0.061	0.045	0.059
N(Students)	920541	920541	916760	916760	920541	920541	920541	920541	920541	920541

Notes: This table replicates the analysis in Table 3 from the main text but using English language arts achievement as the outcome. The notes to Table 3 apply.

Appendix Table A.5. Estimates of the Math Achievement Gap in Grades 3-8 by FRM Coding Status, Pseudo-Coding Scenario-3, but with Random Assignment of Implementation of the CEP across Schools.

	Pre-CEP Pseudo-Coding 3, Random Pseudo-Coding, Equation 1		Pre-CEP Pseudo-Coding 3, Random Pseudo-Coding, Equation 2	
	(1)	(2)	(3)	(4)
FRM	-0.570	-0.398	-0.468	-0.345
	(0.013)***	(0.010)***	(0.006)***	(0.004)***
FRM Aggregate			-0.364	-0.197
			(0.045)***	(0.039)***
Other Controls		Y		Y
Share of Students FRM	56.5%		56.5%	
Share of Schools CEP	30.7%		30.7%	
R-Squared	0.082	0.221	0.088	0.225
N(Students)	916461	916461	916461	916461

Notes: This table replicates the results in columns (9) and (10) of Tables 2 and 3, except schools are randomly assigned as CEP switchers, rather than using CEP eligibility rules. The results can be compared to the results in Tables 2 and 3 to assess the extent to which the concentration of miscoded students at high-poverty schools reduces the impact of the CEP on model performance. The notes to Tables 2 and 3 apply.

Appendix Table A.6. The Effect of CEP Coding on School Accountability Based on Value-Added to Test Scores in English Language Arts, Grades 3-8.

	Pre-CEP	Post-CEP	Pre-CEP with Pseudo-Coding 1
Average percentile ranking in the school distribution of year-1 CEP adopters (in 2014-2015) using VAM for all students	49.3	51.3	55.5
Number of year-1 CEP adopters (in 2014-2015) ranked in the top quintile using VAM for all students	49	49	69

Notes: This table replicates the results reported in Table 4 in the main text, but using value-added to English language arts achievement. The notes to Table 4 apply.

Appendix Table A.7. Estimated English Language Arts Achievement Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17).

	(1)	(2)	(3)	(4)	(5)	(6)
Individual FRM	-0.656		-0.508		-0.405	-0.293
	(0.012)***		(0.006)***		(0.006)***	(0.004)***
Individual DC		-0.591		-0.414	-0.195	-0.170
		(0.010)***		(0.006)***	(0.004)***	(0.003)***
FRM School Share			-0.569		0.159	0.026
			(0.031)***		(0.070)**	(0.074)
DC School Share				-1.285	-1.171	-0.972
				(0.035)***	(0.099)***	(0.110)***
Other Controls						Y
R-Squared	0.115	0.080	0.131	0.123	0.145	0.271
N (Students)	914834	914834	914834	914834	914834	914834

Notes: This table replicates the analysis in Table 5 from the main text but using English language arts achievement as the outcome. The notes to Table 5 apply.

Appendix Table A.8. Estimated English Language Arts Achievement Gaps in Grades 3-8 Using FRM and DC Information During the Post-CEP Period (Years 2015-17), Inclusive of Panel Measures.

	(1)	(2)	(3)	(4)
Panel FRM	-0.641		-0.537	-0.443
	(0.014)***		(0.011)***	(0.009)***
Panel DC		-0.466	-0.109	-0.100
		(0.009)***	(0.007)***	(0.006)***
Contemporary Individual and School-Aggregated FRM	Y		Y	Y
Contemporary Individual and School-Aggregated DC		Y	Y	Y
Other Controls				Y
R-squared	0.151	0.130	0.160	0.281
N (students)	914834	914834	914834	914834

Notes: This table replicates the analysis in Table 7 from the main text but using English language arts achievement as the outcome. The notes to Table 7 apply.

Appendix Table A.9. Replication of Full-Model Math Achievement Results in Table 7 Using Only the Final Year of the Data Panel (2017).

	(1)	(2)
Panel FRM	-0.500	-0.405
	(0.013)***	(0.012)***
Panel DC	-0.126	-0.106
	(0.010)***	(0.009)***
Contemporary Individual and School-Aggregated FRM	Y	Y
Contemporary Individual and School-Aggregated DC	Y	Y
Other Controls		Y
R-squared	0.154	0.255
N (students)	306646	306646

Notes: This table extends the analysis in Table 7 in the main text. The notes to Table 7 apply.