Economics of Education Review 94 (2023) 102374

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Free and reduced-price meal enrollment does not measure student poverty: Evidence and policy significance

Ishtiaque Fazlul [a], Cory Koedel [b,*], Eric Parsons [c]

[a] Department of Economics, Finance, and Quantitative Analysis, Kennesaw State University, United States
[b] Department of Economics & Truman School of Government and Public Affairs, University of Missouri, United States
[c] Department of Economics, University of Missouri, United States

ARTICLE INFO

Keywords:
School poverty
Free and reduced price lunch
Free and reduced price meals
School funding
School accountability
School neighborhood poverty
Direct certification

ABSTRACT

Free and reduced-price meal (FRM) enrollment is commonly used in education research and policy applications as an indicator of student poverty. However, using multiple data sources external to the school system, we show that FRM status is a poor proxy for poverty, with enrollment rates far exceeding what would be expected based on stated income thresholds for program participation. This is true even without accounting for community eligibility for free meals, although community eligibility has exacerbated the problem in recent years. Over the course of showing the limitations of using FRM data to measure poverty, we also provide early evidence on the potential value of two alterative measures of school poverty.

1. Introduction

Free and reduced-price meal (FRM) enrollment under the National School Lunch Program (NSLP) plays a central role in identifying high-poverty students in U.S. education policy. For example, under the federal Every Student Succeeds Act (ESSA), the accountability systems of all fifty states plus Washington DC track gaps in student achievement by poverty status, and 84 percent (all but six states and Washington DC) use FRM data to identify high-poverty students. FRM data are also used to allocate federal, state, and local funding with the goal of targeting resources toward schools serving low-income children.[1] The scholarly community is similarly reliant on FRM data to identify high-poverty students for a variety of research-based applications (Domina et al., 2018).

While FRM data are commonly used in these roles, it is well-understood that FRM designations are error-prone, blunt indicators of poverty that obscure wide variation in income within FRM status bins (Domina et al., 2018; Harwell & LeBeau, 2010; Michelmore & Dynarski, 2017; Parsons, Koedel & Tan, 2019). There is also evidence that FRM status is awarded to more students than income-eligibility thresholds would imply. For example, Bass (2010) shows disparate trends in youth poverty rates measured inside and outside of schools from the 1970s through the early 2000s, while Domina et al. (2018) link FRM data to IRS tax records and show that FRM status is awarded to more students than income data suggest should be eligible.

We complement these previous studies by using two new data sources to evaluate the accuracy of FRM data for measuring student poverty in the state of Missouri. The first is administrative records on student direct certification (DC) status. DC data capture participation in social service programs outside of public schools, for which income-eligibility is more carefully vetted than participation in the NSLP. The second data source is recently-developed "school neighborhood poverty" (SNP) metrics made available by the National Center for Education Statistics (NCES). These metrics are estimated for schools using data on incomes of nearby households and were first made available by the NCES in 2016. Both data sources are independently promising for measuring poverty. We increase our confidence in their reliability by validating them against each other, which confirms they contain similar information (on average) in Missouri. We then use them to assess the accuracy of NSLP-based poverty designations.

Our findings complement previous work by Bass (2010) and Domina et al. (2018) by showing that students' FRM designations overstate poverty. We extend the literature by directly estimating the magnitude of the overstatement, which is substantial—for instance, we find that

* Corresponding author.
E-mail address: koedelc@missouri.edu (C. Koedel).

[1] According to data compiled by EdBuild, 33 states use FRM data to allocate increased funding toward students from low-income households (e.g., see here, retrieved on 04.11.2021: http://funded.edbuild.org/reports).

https://doi.org/10.1016/j.econedurev.2023.102374
Received 24 May 2021; Received in revised form 25 January 2023; Accepted 13 February 2023
Available online 10 March 2023
0272-7757/

free-meal enrollment in Missouri is oversubscribed by 35–50 percent relative to the stated income threshold in the NSLP. We further show that NSLP enrollment was oversubscribed prior to the NSLP's community eligibility provision (CEP), though the CEP has made it worse.

Our most credible estimates of oversubscription in the NSLP apply to Missouri and the income threshold at 130 percent of the poverty line, where we can triangulate all three of our data sources: NSLP data, DC data, and SNP data. The DC data are a limiting factor to expanding our analysis: we only have these data from Missouri, and they only plausibly identify children from families at this specific income threshold. However, under some additional assumptions (described below), we extend our investigation of students' NSLP designations using the SNP data in two ways. First, we examine free-meal (FM) and reduced-meal (RM) designations separately. Our findings suggest most of the oversubscription in the NSLP is in the "free meal" category. Second, we analyze a larger sample of 27 states and estimate that NSLP enrollment is oversubscribed in the larger sample at a rate similar to Missouri, on average. The 27-state expansion further reveals considerable heterogeneity across states in the mapping between SNP- and NSLP-based poverty measures, which we identify as an important area for exploration in future research.

We contribute to a thin literature on a topic of great importance for contemporary education research and policy in the United States. We show that FRM data do not measure poverty in public schools accurately and that the errors are substantial and asymmetric. There are several possible explanations for the persistent overstatement of poverty rates in FRM data, with a prominent one being that districts are incentivized to identify students as FRM-eligible but are not similarly incentivized to do so accurately.

These findings refute the common misperception that FRM data have historically served to measure poverty (e.g., prior to the CEP). Instead, as our results make clear, FRM data have captured the nebulous concept of student disadvantage, albeit under the guise of measuring poverty. This longstanding misperception is causing two problems in contemporary education policy. First, it is hampering the use of modern data systems to develop new and more accurate poverty metrics. For instance, old rates of FRM enrollment are being used as benchmarks for assessing the accuracy of new poverty metrics (e.g., see Croninger, Rice & Checovich, 2015; Grich, 2019; Massachusetts Department of Elementary & Secondary Education, 2017). In addition, in some states, CEP-adopting schools and districts that no longer collect FRM data are being forced to report poverty using alternative metrics. If these schools and districts do not build measures that match (high) FRM-based poverty rates, it can put them at a disadvantage in state funding and other policies, even if the FRM-based rates are not correct.[2] Second, it contributes to resistance to the idea of moving away from poverty designations in policy applications and toward the arguably more useful concept of disadvantage designations. Hence, a greater awareness of what FRM data really measure may lead to more productive conversations about measuring student need in public schools.

2. Background

FRM enrollment for individual students under the NSLP is determined by school districts. Districts assess eligibility in two ways. First, students can be "directly certified" for free meals if they participate in a qualified federal assistance program such as the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), or the Food Distribution Program on Indian Reservations (FDPIR). In addition, foster, migrant, runaway, and homeless youth—and in some states, additional income groups like students eligible for Medicaid—can also be directly certified (Gindling et al., 2018; Greenberg, 2018, 2019).

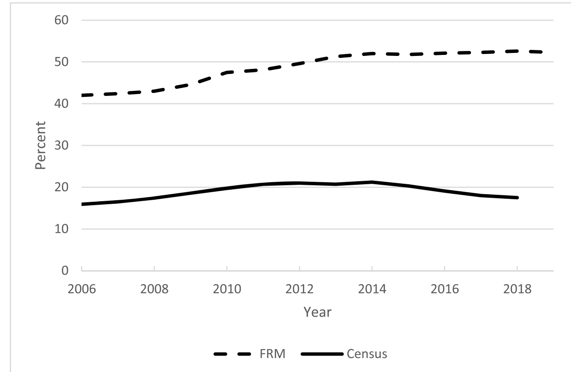

Fig. 1. Poverty Rates Among School-Aged Children Measured Using Different Data Sources.
Notes: Trends are in the share of FRM-eligible students and the share of school-aged children living in families with incomes at or below the poverty line. Data Source: NCES Digest of Education Statistics (de Brey et al., 2021).

Second, school districts administer income surveys to parents, and students can be classified as eligible for free or reduced-price meals based on the survey responses. Students from families with incomes at or below 130 percent of the federal poverty line are eligible for free meals, and those from families with incomes between 130 and 185 percent of the poverty line are eligible for reduced-priced meals. In addition to concerns about student welfare, districts are incentivized to encourage and approve parent applications because they receive meal subsidies for FRM-eligible students and can gain access to additional federal, state, and local funding. Parents' incentives are also aligned—they benefit because participation in the NSLP lowers the cost of food for their children.

Only a small fraction of NSLP applications go through an income verification process (Bass, 2010).[3] In fact, according to the USDA's Eligibility Manual for School Meals in 2017, attempting to verify more than three percent of applications without special cause is prohibited.[4] In instances where income-eligibility cannot be verified—which is quite common, up to 50 percent of applications (Burghardt, Silva & Hulsey, 2004)—FRM status is canceled, but there are no other repercussions. As a result, the incentive structure clearly favors districts and parents stretching the boundaries of eligibility. We do not take a normative stance on whether this is good or bad from a policy perspective and note that research shows providing students with free meals leads to better test scores (Ruffini, 2021; Schwartz & Rothbart, 2020), improved student discipline (Gordon & Ruffini, 2021), and higher wages later in life (Lundborg, Rooth, & Alex-Petersen, 2022).[5] But for the purpose of relying on FRM data to measure poverty, the incentive structure is cause for concern.

Fig. 1 updates a similar figure in Bass (2010) using data from the Digest of Education Statistics through 2018 (de Brey et al., 2021). It

[2] Gindling et al. (2018) provides a useful case study in Baltimore City Public Schools.

[3] See NSLP Verification Toolkit from USDA. Retrieved on 03.30.2021 from https://www.fns.usda.gov/cn/verification-toolkit

[4] The 2017 version of USDA's Eligibility Manual for School Meals says, "With the exception of verification for cause, LEAs must not verify more or less than the standard sample size or the alternate sample size (when the alternate sample size is used). LEAs must not verify all (100 percent) of the applications." Verification for cause can be performed if "the LEA is aware of additional income or persons in the household." This information was retrieved from the following address on 03.31.2021. https://fns-prod.azureedge.net/sites/default/files/cn/SP36_CACFP15_SFSP11-2017a1.pdf

[5] There is also no evidence of increases in BMI or the probability of being obese or overweight (Davis & Musaddiq, 2019; Schwartz & Rothbart, 2020).

Table 1
Summary statistics, Missouri data.

	2016		2017	
	Mean	Standard Deviation	Mean	Standard Deviation
Demographics				
Black	0.16	0.26	0.16	0.26
Hispanic	0.06	0.08	0.06	0.09
White	0.72	0.28	0.71	0.28
Multi-race	0.03	0.03	0.04	0.03
Asian/Indian/Pacific Islander	0.03	0.03	0.03	0.04
Female	0.49	0.03	0.49	0.03
IEP	0.13	0.08	0.14	0.08
ESL	0.04	0.08	0.05	0.09
Test Scores				
Standardized Math Score	0.00	0.45	0.00	0.44
Poverty Measures				
Share Free/Reduced-Price Meal Eligible	0.53	0.26	0.52	0.26
Share Free Meal Eligible	0.47	0.27	0.46	0.27
Share of Directly Certified	0.30	0.18	0.30	0.18
NCES IPR Estimate	284.11	137.88	289.84	138.81
IPR(130)	0.34	0.11	0.33	0.11
IPR(185)	0.42	0.13	0.41	0.13
Avg. Students Per School	423.47	341.14	421.47	340.71
N (Schools)	2,172		2,186	
N (Students)	919,786		921,335	

Notes: This table shows summary statistics for our analytic sample of schools in Missouri in the 2016 and 2017 school years with at least 25 students. The summary statistics are weighted by enrollment. Student demographics, test scores, free and reduced-price meal enrollment, and direct certification status are taken from Missouri administrative microdata. IPR estimates are taken from the NCES school neighborhood poverty (SNP) metrics. IPR(130) and IPR(185) are calculated from the reported IPR estimates and standard errors for each school as described in the text. Test scores are from a reduced sample of schools that have test-takers in grades 4–8. The test-taking school samples from 2016 to 2017 include 1689 and 1694 schools, respectively.
Data Source: DESE administrative data and SNP data from NCES, 2016 and 2017.

plots the national share of FRM-eligible students and the share of school-aged children living at or below the poverty line. The former data are collected by school districts as described above; the latter are based on data from the U.S. Census.[6] The income thresholds corresponding to these poverty definitions are different—i.e., the stated FRM enrollment threshold is at 185 percent of the poverty line—which limits comparative inference to some degree. Still, the differential trends in the two poverty measures over the 2006–2018 period suggest a possible measurement problem. Most notably, whereas the share of children in poverty according to the Census moves with the business cycle as anticipated and increases by just 1.5 percentage points from 2006 to 2018, the FRM-eligible share rises throughout the sample period and increases by more than 10 percentage points.

The most closely-related study to our own is Domina et al. (2018), who merge FRM enrollment data from Oregon, and a single school district in California, with family income data from the IRS. These authors find disagreement in the data in both directions—i.e., seemingly FRM-eligible students based on income who are not enrolled in the NSLP and income-ineligible students who are enrolled. However, consistent with our findings below, the latter are more prevalent than the former. Moreover, the data used by Domina et al. (2018) predate the CEP, which is a provision of the NSLP that allows schools and districts to provide free meals to all students if the student body is sufficiently impoverished. In many states, FRM data for CEP schools are overwritten to indicate that 100 percent of students are FM-enrolled (Chingos, 2016; Greenberg, Blagg & Rainer, 2019; Koedel & Parsons, 2021). As a result, the CEP further degrades the link between student poverty and FRM enrollment. In the analysis that follows, we show that the CEP contributes to the overstatement of poverty in modern FRM data, but it is not the primary driver and even in the absence of the CEP, FRM data still greatly overstate poverty.[7]

3. Data

3.1. Missouri administrative data

We conduct our primary analysis using administrative student records from the Missouri Department of Elementary and Secondary Education (DESE) for students enrolled in public schools during the 2015–16 and 2016–17 school years (school years are hereafter identified by the spring year—e.g., 2016 for 2015–16). We restrict our analysis to schools with at least 25 students. The most important variables in the administrative data are students' "free" and "reduced-price" meal designations (FM and RM, respectively). Through DESE, we also have access to merged administrative data indicating whether each student is directly certified to receive free meals. We refer to this combined dataset as the Missouri administrative data. Table 1 provides descriptive information about our sample.

Students from households that participate in SNAP, TANF, and FDPIR, and students classified as migrant, runaway, homeless, or in foster care are categorically eligible for free meals in Missouri.[8] DESE has an agreement with the Missouri Department of Health and Senior

[6] These data are reported across several issues of the Digest of Education Statistics, the most recent of which is de Bray et al. (2021).

[7] One might worry about the impact of the CEP on the FRM trend in Figure 1, but no impact is visually apparent. There are two reasons: (1) some states have not overwritten their FRM enrollment data, dulling its impact nationally, and (2) even in states where the CEP has overwritten the data, only a small fraction of the total student population is affected despite large changes in some schools and districts (Koedel & Parsons, 2021).

[8] Information retrieved from the following address on 04.04.2022:https://dese.mo.gov/media/pdf/free-and-reduced-application-and-direct-certification-information-and-procedures-2021.

Services to provide the program-participation information necessary to directly certify these students. All Missouri districts are required to download direct certification information for their students at least three times annually to ensure all students eligible for FM through direct certification are extended the benefit. Missouri's direct certification processes are above-average among states along several measurable dimensions (Koedel & Parsons, 2021).

The key feature of the direct-certification landscape in Missouri that facilitates our analysis is that these criteria should identify students living at or below 130 percent of the poverty line, which is the same threshold used by the NSLP to determine FM eligibility. This is because SNAP is the primary program through which students are directly certified and it uses this threshold. In most other states, the DC and FM eligibility thresholds are not aligned because broad-based categorical eligibility (BBCE) policies allow families with higher incomes to qualify for SNAP.[9] The income-threshold alignment between direct certification and FM eligibility in Missouri makes it an ideal setting in which to use DC data to assess the accuracy of NSLP data. Below we discuss the generalizability of our findings outside of Missouri given the somewhat unique circumstance that Missouri lacks BBCE (Missouri is one of just six states without BBCE).

Although by the intent of the rules direct certification in Missouri should identify students at or below 130 percent of the poverty line, this may not happen in practice. Formally, we require the following assumptions to hold in order for the share of students who are directly certified in our data to reflect the share of students living at or below 130 percent of the poverty line:

1. Income-eligibility requirements for the social-service programs that lead to direct certification include students up to 130 percent of the poverty line and are strictly enforced.
2. All eligible families participate in social-service programs that lead to direct certification.

It seems implausible that these assumptions are never violated, but there are reasons to expect they hold, at least to a close approximation. The absence of BBCE in Missouri simplifies income-eligibility rules, making it more likely they are enforced as stated (Blankley, 2019). Moreover, the national participation rate in SNAP among families with children is estimated to be 100 percent (Schanzenbach, 2019), and Missouri's total participation rate in SNAP (inclusive of all eligible individuals) is well above the national average.[10]

3.2. NCES school neighborhood poverty data

Beginning in 2016, the NCES began reporting SNP metrics for nearly every school in the United States.[11] These metrics are based on household income data from the U.S. Census Bureau's American Community Survey (ACS) and are reported as continuous variables that measure the average income-to-poverty ratio (IPR) in a school, multiplied by 100. For example, a value of exactly 100 indicates the average income is at the poverty line, a value of 200 indicates the average income is double the poverty line, and so on. The IPR metrics are described in Geverdt (2019) as capturing "economic conditions of neighborhoods where schools are located," but to be more precise, they capture the income-to-poverty ratio in a household that would hypothetically be situated in the exact geographic location of the school.

We elaborate briefly on the construction of the SNP metrics here and refer interested readers to Geverdt (2019) for more information. The SNP metrics are estimated using a spatial estimation process called Kriging. This method uses the weighted sum of income values in measured locations to predict values in unmeasured locations (Cressie 1989; Cressie 1993). The predicted value in the unmeasured location is estimated by the following equation (Geverdt & Nixon, 2018):

$$\widehat{Z}(s_0) = \sum_{i=1}^{\mathbb{N}} \lambda_i Z(s_i) \tag{1}$$

where $\widehat{Z}(s_0)$ is the predicted income value in the unmeasured location, $Z(s_i)$ is the value at measured location i, and λ_i is a weighting parameter. The closer that measured location i is to the unmeasured location, the larger is λ_i. The value of λ_i also depends on the covariance structure of all measured locations; i.e., the relationship between distance and income elsewhere in the data, which is modelled using a semi-variogram when calculating SNP (Geverdt & Nixon, 2018). The NCES IPR estimates for each school are based on data from the 25 households closest to the school in the American Community Survey.[12]

While conceptually compelling, whether IPR is an accurate measure of school poverty is uncertain, and we are not aware of any prior work validating its accuracy. Under the following assumptions, IPR will be an accurate indicator of poverty in a school:

1. It must be the case that school enrollment is comprised of students who live in the area surrounding the school. Broadly speaking, private school enrollment and school choice programs that alter geographically proximal enrollment are threats to this assumption.
2. Estimated poverty at the exact location of a school reflects poverty in a school's catchment area. If homes closer or farther from schools within catchment areas have systematically higher incomes, this assumption would be violated.

Again, it is implausible that these assumptions are never violated. However, whether violations are common or cause systematic bias is uncertain.

In addition, to facilitate our comparison to the FM (and DC) data, we must manipulate the continuously-measured IPR values to estimate the share of students living at or below 130 percent of the poverty line for each school. Our manipulation of the IPR values relies on an additional assumption that they are mean values from a normal distribution. Under this assumption, we use the IPR estimates and their standard errors (also reported by NCES), which we convert to standard deviations by multiplying them by $\sqrt{N}$, to construct the distribution of income in each school. Then, the fraction of students with incomes at or below any threshold value can be calculated directly from the cumulative distribution function (CDF). Eq. (2) gives an example at the focal value of 130 percent of the poverty line:

[9] As of January 2022, the gross income limit for BBCE across states ranged from 130-200 percent of the poverty line, with most states falling in the upper end of this range (see United States Department of Agriculture, 2022).

[10] SNAP participation rates for U.S. states can be found here: https://www.fns.usda.gov/usamap# (retrieved 07.01.2022). A notable contextual feature of Missouri is the small Hispanic population share. Research suggests that assumption 2 is more likely to be violated in states/locales with large Hispanic populations (Lichter et al., 2015; Sandstrom et al., 2014; Williams, 2013; Zedlewski & Martinez-Schiferl, 2010).

[11] For example, in Missouri, 2,172 out of the 2,215 public schools have corresponding SNP metrics from NCES in 2016 (98 percent). In 2017, SNP metrics are available for 2,186 out of 2,219 schools (99 percent).

[12] An issue with basic Kriging is that it assumes the relationship between the variance of the measure and distance between locations is the same throughout the sample. But that may not be true for SNP estimation conducted on a national scale covering a variety of regions and regional contexts (e.g., urban versus rural areas). NCES's SNP metrics are estimated using empirical Bayesian Kriging, which addresses this problem by dividing areas into smaller regions and developing models for each region (Geverdt & Nixon, 2018). The local models take into account differences in spatial dependence across regions.

$$\widehat{IPR(130)}_{jt} = P\left(IPR_{jt} \leq 130\right) = \int_{-\infty}^{130} f(IPR)dIPR \tag{2}$$

In the equation, $\widehat{IPR(130)}_{jt}$ is the estimated fraction of students in school j and year t with family incomes at or below 130 percent of the poverty line, and $f(IPR)$ is the probability density function of IPR. The general form of Eq. (2), where X indicates a generic income value as a percent of the poverty line, can be written as:

$$\widehat{IPR(X)}_{jt} = P\left(IPR_{jt} \leq X\right) = \int_{-\infty}^{X} f(IPR)dIPR \tag{3}$$

This procedure for manipulating the underlying IPR values adds to the list of assumptions under which our SNP-based poverty metrics, IPR (X), are accurate measures for individual schools. The normality assumption is useful but an approximation. An obvious technical limitation is that normality allows for negative income values. However, this limitation is of little concern in our application because we only care about the area under the curve below a certain threshold (either 130 or 185 percent of the poverty line). Of greater importance is whether the shapes of schools' local-area income distributions are approximately normal around the IPR estimates. While we lack data to make a conclusive statement in this regard, aspects of our estimates are consistent with the normality assumption working well empirically. Moreover, the normality assumption fits with the estimation procedure used by NCES to produce the initial IPR values.[13]

In addition to the normality assumption, our conversion of the standard errors to standard deviations assumes that the original IPR values are unweighted, but in reality they are weighted averages. To elaborate briefly, we multiply the standard error of each IPR estimate by $\sqrt{N}$ to get the standard deviation, but this is only the correct conversion if the data are unweighted. With weighted data, the multiplicative factor is variable, and its average value will be smaller than $\sqrt{N}$ depending on the variance of the weights within a school (the more variance, the smaller the value). Unfortunately, the weights are not available from NCES.[14] Conceptually, there is good reason to expect the variance of the weights within a school to be modest (e.g., residential sorting by income), in which case our $\sqrt{N}$ simplification is reasonable. Moreover, there are other aspects of the NCES estimation process that may work in the opposite direction, in which case our "too large" $\sqrt{N}$ adjustment may be offsetting, on average.[15]

Compared to our administrative DC data, there is more *ex ante* uncertainy about the accuracy of IPR(X). Still, we view the assumptions under which we recover IPR(X) as reasonable, and we only require IPR (X) to be accurate on average to be useful in our anlaysis (see below for details). The bottom panel of Table 1 provides basic summary statistics for IPR as reported by NCES, along with our modified versions of IPR—IPR(130) and IPR(185).

4. Methods

We begin by comparing school shares of students living at or below 130 percent of the poverty line, as estimated using the DC and SNP data. The assumptions under which each estimate is accurate, documented in the preceding section, are very different. If either or both sets of assumptions are violated, a divergence of the estimates seems almost assured. But if the estimates agree, it is difficult to construct a story by which they are both wrong but the sources of errors just happen to align such that they are wrong in the same direction and to the same degree.

Following this logic, we compare estimates of $IPR(130)$ to schools' DC shares using the following univariate regression, weighted by school enrollment:

$$DC_{jt} = \beta_0 + \widehat{IPR(130)}_{jt}\beta_1 + \varepsilon_{jt} \tag{4}$$

In Eq. (4), DC_{jt} is the share of directly-certified students in school j in year t and $\widehat{IPR(130)}_{jt}$ is the estimated value from Eq. (2). If both variables in this regression are measuring the same construct, on average, then the expected value of β_1 is 1.0. Deviations from 1.0 would imply systematic differences in what they measure. Note that the empirical Bayesian Kriging procedure used to construct the original IPR variables embeds shrinkage, so attenuation bias in β_1 is not a concern (Chetty, Friedman & Rockoff, 2014; Jacob & Lefgren, 2008).

A sufficient (albeit not necessary) condition for recovering a value of $\beta_1 = 1.0$ is that the assumptions outlined above for each measure are satisfied; or, at least satisfied to a rough approximation. Below we show that we fail to reject the null hypothesis that $\beta_1 = 1.0$ in Eq. (4) with a fairly precise confidence interval. This implies that both the DC share and IPR(130) are accurate indicators of the share of students living at or below 130 percent of the poverty line in Missouri, at least on average. Taking this as a point of departure, we then estimate the following univariate regressions, also weighted by school enrollment:

$$FM_{jt} = \gamma_0 + DC_{jt}\gamma_1 + u_{jt} \tag{5}$$

$$FM_{jt} = \delta_0 + \widehat{IPR(130)}_{jt}\delta_1 + e_{jt} \tag{6}$$

In Eqs. (5) and (6), we regress the share of students eligible for FM in school j and year t, FM_{jt}, on the school's DC share and IPR(130) estimate, respectively. By rule, students identified as eligible for FM should include only those in households at or below 130 percent of the poverty line. Therefore, the same logic from Eq. (4) applies—we should anticipate that δ_1 and γ_1 have values of 1.0. Values above 1.0 would indicate that more students are designated for free-meal status than income-eligibility alone would dictate.[16]

We also extend Eq. (6) to look at the threshold for free *and reduced-price* meal enrollment, which is at 185 percent of the poverty line, using Eq. (7):

$$FRM_{jt} = \lambda_0 + \widehat{IPR(185)}_{jt}\lambda_1 + \eta_{jt} \tag{7}$$

In Eq. (7), FRM_{jt} is the share of students eligible for free or reduced-price meals, and $\widehat{IPR(185)}$ is the income-aligned measure based on the SNP data. λ_1 takes on the same interpretation as δ_1 and γ_1 above—i.e., values above 1.0 indicate oversubscription for free and reduced-price meals. A limitation of this extension is that while it is motivated by the validation regression in Eq. (4) at 130 percent of the poverty line, we do not have any external data to assess the validity of $IPR(185)$. We must assume that our findings for the comparison of $IPR(130)$ and the DC share imply that $IPR(185)$ is also an accurate measure of the fraction of

[13] In estimating the IPR values from the underlying household income data, NCES makes a transformation to bring the data closer to a normal distribution (Geverdt & Nixon, 2018). This likely contributes to our effective estimation of IPR(X) under the normality assumption. In addition, compared to what is arguably the most reasonable alternative distributional assumption—lognormality—our estimates that assume normality triangulate better with existing information about poverty and social program participation (Schanzenbach, 2019) and NSLP participation (Domina et al., 2018).

[14] Information on the weights is necessary to make the technically correct conversion, but this information is not reported by NCES. Moreover, correspondence with NCES indicates they do not have this information and it was not covered in their release agreement with the Census, making it infeasible to obtain (at least in the near term).

[15] In particular, IPR is an estimate of household income at the precise location of the school, not the average income of students in the area, and it is reasonable to expect the variance of precise-location income to be lower than local-area average income, all else equal.

[16] There is some nuance to this interpretation—see Domina et al. (2018) for a discussion.

Table 2
Univariate alignment regressions, Missouri data.

VARIABLES	(1) 2016 Dependent variable: DC share	(2) 2017 Dependent variable: DC share	(3) 2016 Dependent variable: FM share	(4) 2017 Dependent variable: FM share	(5) 2016 Dependent variable: FM share	(6) 2017 Dependent variable: FM share	(7) 2016 Dependent variable: FRM share	(8) 2017 Dependent variable: FRM share
IPR(130)	1.026	0.994			1.505***	1.469***		
	(0.033)	(0.034)			(0.048)	(0.051)		
DC share			1.372***	1.386***				
			(0.014)	(0.015)				
IPR(185)							1.385***	1.396***
							(0.034)	(0.037)
Constant	−0.045†††	−0.036†††	0.049†††	0.050†††	−0.046†††	−0.030†	−0.052†††	−0.050†††
	(0.011)	(0.011)	(0.004)	(0.004)	(0.016)	(0.017)	(0.015)	(0.016)
Observations	2,172	2,186	2,172	2,186	2,172	2,186	2,172	2,186
R-squared	0.370	0.348	0.849	0.830	0.359	0.328	0.468	0.440

Notes: This table presents estimates from school level univariate regressions weighted by enrollment in each school year. In each regression, we test the null hypothesis that the poverty-measure coefficient is 1.0; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For completeness, we also report on the statistical significance of the constant term, where †††, ††, and † indicate the constant is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
Data Source: DESE administrative data and SNP data from NCES, 2016 and 2017.

families living at or below 185 percent of the poverty line, on average. While reasonable, we have no way of providing direct evidence to confirm or refute this assumption, and thus we present our findings from Eq. (7) as suggestive only.

Finally, we return to the point above that like many other states, the FRM data in Missouri are affected by the CEP. This means some high-poverty schools are coded as entirely comprised of FM students even when individual income-eligibility is below 100 percent. One could interpret the CEP as "biasing" upward the estimates in Eqs. (6) and (7), although in our view the term "bias" is not appropriate because the CEP is a true source of inaccuracy in modern FRM data. Still, we assess the impact of the CEP on the estimates in Eqs. (5), (6), and (7) by imputing the FM_{jt} and FRM_{jt} values for CEP schools to their values during the last year prior to CEP implementation in Missouri, which was 2014. If the CEP was solely responsible for the oversubscription in FRM data, we would expect our estimates using the CEP-adjusted data to yield coefficients on the key parameter of 1.0.

5. Free and reduced-price meal enrollment does not measure student poverty

5.1. Primary findings

Table 2 shows results from our baseline regressions in Eqs. (4), (5), (6), and (7). The column headers indicate the dependent variable in each model and the year for which the model is estimated (either 2016 or 2017, which are the first two years SNP metrics were published by NCES).

First, columns (1) and (2) report results from the alignment regressions of the DC share on IPR(130). We cannot reject the null hypothesis that $\beta_1 = 1.0$, and our confidence intervals are precise. This is consistent with the aforementioned assumptions being upheld under which these two measures converge, at least on average.

Next, columns (3)-(6) show regressions of the FM share on the DC share and IPR(130), respectively, as shown in Eqs. (5) and (6). If we believe that FM enrollment follows the income-eligibility rules, we should also get coefficients of 1.0 in these regressions, but our estimates are much larger. The coefficients range from 1.37 to 1.51, implying an oversubscription rate for FM in the range of 37–51 percent. In all cases in columns (3)-(6), we can comfortably reject the null hypothesis of a 1.0 coefficient.

We make two additional observations about these estimates. First, the coefficients on the DC-share variables are somewhat smaller than on IPR(130). We do not explore this result in depth, but note that modest differences along the lines of what we find are not ruled out by the results in columns (1) and (2). This is because those results show that the DC share and IPR(130) provide the same information about poverty on average, but there are distributional differences in the variables that could contribute to differences in the coefficients in columns (3)-(6). Second, the standard errors in the IPR(130) regressions are much larger while the R-squared values are smaller, reflecting greater imprecision in these estimates relative to the DC shares based on the Missouri administrative data. We elaborate on these issues below and in Appendix A.

In the last two columns of Table 2 we present the comparisons between the FRM enrollment share and IPR(185). These results also indicate NSLP oversubscription, with coefficients in 2016 and 2017 of 1.385 and 1.396, respectively. These estimates are most comparable to the IPR (130) estimates in columns (5) and (6) because they use the same measurement mode. Inference based on both sets of estimates suggests the FM oversubscription rate exceeds the RM meal oversubscription rate. The lower oversubscription rate in RM data is alluded to in our descriptive statistics in Table 1, which show few students are listed as eligible for reduced-price meals.

Next, we consider the possibility that the CEP is driving the oversubscription of FM and FRM enrollment in our data. We build the modified dataset described above in which the FM and FRM shares for non-CEP schools are left as reported in 2016 and 2017, but for CEP schools, these values are re-coded to the last pre-CEP year in Missouri: 2014. Table 3 shows the results. The coefficients in Table 3 decline in both the FM and FRM regressions compared to their analogs in Table 2, but still imply substantial and statistically significant oversubscription in the NSLP. Indeed, Table 3 shows the majority of the oversubscription in the NSLP indicated by Table 2 is not due to the CEP (the reduced coefficients from the FM (FRM) models in Table 3 indicate that the CEP accounts for about 15 (9) percentage points of the total oversubscription rates estimated Table 2).[17]

5.2. Supplementary analysis: student poverty and student achievement

We also use the poverty metrics to predict student achievement using the following cross-sectional regression at the school level:

[17] In results suppressed for brevity, we confirm our findings in Table 3 are insensitive to using a more complex imputation procedure for CEP schools that accounts for general trends in FM and FRM enrollment in Missouri (among non-CEP schools). The reason for the insensitivity is that similarly to the national data (Figure 1), FM and FRM enrollment trends in Missouri are essentially flat from 2014-2017.

Table 3
Univariate alignment regression with imputed FM and FRM share for CEP schools, Missouri data.

VARIABLES	(1) 2016 Dependent Variable FM Share	(2) 2017 Dependent Variable FM Share	(3) 2016 Dependent Variable FM Share	(4) 2017 Dependent Variable FM Share	(5) 2016 Dependent Variable FRM Share	(6) 2017 Dependent Variable FRM Share
DC Share	1.227***	1.235***				
	(0.009)	(0.010)				
IPR(130)			1.354***	1.318***		
			(0.040)	(0.043)		
			0.000	0.000		
IPR(185)					1.293***	1.308***
					(0.030)	(0.032)
Constant	0.069†††	0.069†††	−0.020	−0.008	−0.033††	−0.035††
	(0.003)	(0.003)	(0.014)	(0.014)	(0.013)	(0.014)
Observations	2,160	2,168	2,160	2,168	2,160	2,168
R-squared	0.923	0.913	0.398	0.373	0.504	0.490

Notes: This table presents estimates from school level univariate regressions weighted by enrollment in each school year. For CEP schools, the FM and FRM shares are imputed to the 2014 level, the last year of non-CEP coded data in Missouri. In results suppressed for brevity, we also confirm these findings are essentially unaffected if we incorporate trends in FM and FRM enrollment in Missouri (based on data from non-CEP schools) into the imputation procedure. In each regression, we test the null hypothesis that the poverty-measure coefficient is 1.0; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For completeness, we also report on the statistical significance of the constant term, where †††, ††, and † indicate the constant is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
Data Source: DESE administrative data, 2016 and 2017, with imputed FM and FRM data from 2014 DESE administrative data for selected schools; and SNP data from NCES, 2016 and 2017.

Table 4
Relationships between student test scores and measured poverty, Missouri data.

VARIABLES	(1) 2017 Dependent Variable: School Avg. Test Score	(2) 2017 Dependent Variable: School Avg. Test Score	(3) 2017 Dependent Variable: School Avg. Test Score
DC Share	−1.685***		
	(0.056)		
IPR(130)		−1.732***	
		(0.167)	
FM share			−1.087***
			(0.043)
Constant	0.535†††	0.577†††	0.527†††
	(0.020)	(0.058)	(0.021)
Observations	1,694	1,694	1,694
R-squared	0.490	0.172	0.454

Notes: This table presents estimates from school-level univariate regressions for 2017 where the dependent variable is the school average standardized math test score, and the independent variables are three different measures of poverty–the DC share, IPR(130), and the FM share in the school. All regressions are weighted by enrollment. In each regression, we test the null hypothesis that the poverty-measure coefficient is zero; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For presentational consistency, we continue to denote statistical significance of the constant term at the 1, 5, and 10 percent levels using the same †††, ††, and † indicators from previous tables. Standard errors are in parentheses.
Data Source: DESE administrative data and SNP data from NCES, 2017.

$$Y_j = \phi_0 + P_j\phi_1 + \varepsilon_j \tag{8}$$

In Eq. (8), Y_j is the average standardized math test score for students in school j, and P_j is a measure of the share of students at 130 percent of the poverty line or below. We estimate this regression three times, where P_j represents either the DC share, IPR(130), or the FM share. The sample includes all schools in Missouri with at least one grade in the 4–8 range (and we continue to impose the condition that enrollment is at or above 25).

It is well-documented that student poverty—despite its imprecise measurement using FRM data—is a strong predictor of low achievement (Domina et al., 2018; Michelmore & Dynarski, 2017). As such, the purpose of Eq. (8) is to compare the implied change in test scores associated with a one-percentage-point increase in the poverty share of a school, as measured by each construct. We expect ϕ_1 to be negative in each version of Eq. (8), and if all three metrics are capturing the same information, it should be similar in magnitude in each regression as well.

Consistent with the preceding analysis, the results in Table 4 suggest the FM share is a less acute measure of poverty than the DC share or IPR (130), which continue to track each other closely (the table shows results from 2017; results from 2016 are similar and omitted for brevity). Specifically, a one-percentage-point increase in the DC share, or IPR (130), corresponds to a reduction in test scores of about 0.017 student standard deviations (note that all poverty variables in Table 4 are on a 0–1 scale). In contrast, the same one-percentage-point increase in the FM share corresponds to a much smaller reduction in student achievement—just 0.011 standard deviations.[18]

A caveat to these results is that the models using IPR(130) are noisier than the other models, which is a general condition present throughout our study. The practical implication is that despite the nominal alignment between the DC-share and IPR(130) coefficients across equations, we cannot rule out fairly large differences in their values with great confidence. Still, our best estimates (i.e., the point estimates) indicate the DC-share and IPR(130) coefficients align closely in terms of their

[18] While the findings in Table 4 are intuitive and follow from the preceding analysis, they are seemingly at odds with results from Domina et al. (2018). Domina et al. (2018) run a series of student-level regressions of test scores on poverty as measured by (a) income tax data from the Internal Revenue Service and (b) FRM status. The tax data allow them to construct more accurate poverty variables, analogously to DC status or IPR(130) here. Domina et al. (2018) generally find larger coefficients on the FRM status variables, from which they conclude, "Perhaps surprisingly, the results of these analyses indicate that school-reported [FRM] status variables are more closely associated with student achievement on standardized tests…than parallel categories constructed using IRS-reported household income" (page 543). In results suppressed for brevity (and available upon request), we replicate their findings substantively using student-level Missouri data. However, we disagree with their interpretation, and at least in Missouri, inference from some of the regressions is confounded by Simpson's paradox (Simpson, 1951). Our own investigation of this issue buttresses our finding that more-acutely measured poverty more strongly predicts low test performance.

Table 5
Univariate alignment regressions, Missouri and the 27-state extended sample using the common core of data.

VARIABLES	(1) 2017 Dependent Variable: School FM Share in MO	(2) 2017 Dependent Variable: School FRM Share in MO	(3) 2017 Dependent Variable: School FM Share in 27 States, Not CEP-Coded	(4) 2017 Dependent Variable: School FRM Share in 27 States, Not CEP-Coded
IPR(130)	1.419***		1.397***	
	(0.052)		(0.008)	
IPR(185)		1.363***		1.253***
		(0.040)		(0.007)
Constant	−0.030†	−0.050†††	−0.045†††	−0.025†††
	(0.017)	(0.017)	(0.006)	(0.006)
State Fixed Effects	N/A	N/A	Yes	Yes
Observations	2,257	2,257	61,270	61,270
R-squared	0.306	0.417	0.477	0.525

Notes: This table presents estimates from school-level univariate regressions weighted by enrollment in each school in 2017 using CCD and SNP data from NCES. In each regression, we test the null hypothesis that the poverty-measure coefficient is 1.0; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For completeness, we also report on the statistical significance of the constant term, where †††, ††, and † indicate the constant is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
Data Source: Common Core of Data and SNP data, both from NCES, 2017.

relationship with academic performance and that both are quite different from the FM coefficient, as in Table 2.[19]

6. Extensions beyond Missouri

In this section, we expand our analysis of the accuracy of students' NSLP designations outside of Missouri leveraging the SNP data. First, we construct IPR(130) and IPR(185) estimates for all schools in the U.S. using the baseline IPR values published by the NCES. Then, we merge these variables with FM and FRM enrollment shares from the Common Core of Data CCD.[20]

There are two challenges associated with this expansion. First, although our analysis in Missouri suggests that IPR(130) is accurate on average in one state, we do not have access to credible validating data in other states to assess its accuracy elsewhere. Therefore, we must assume our findings in Missouri imply that IPR(130) will be accurate, on average, in other states. (We attempted to replicate our tests of the alignment between the DC share on IPR(130) in other states using the CCD. However, we could not confirm the reliability of the DC data in the CCD to support these tests. See Appendix B for details.) And like in our preceding analysis in Missouri, our investigation using IPR(185) must continue to assume our findings for IPR(130) apply to this other income threshold.

The second challenge is with respect to the FM and FRM data in the CCD. Some states have changed how they report these categories due to the CEP and others have not, and there is no indicator in the data to distinguish them. It would cloud inference to evaluate a mix of states coding their data differently. To address this problem, we identify a subset of 27 states that do not appear to have manipulated their FRM reporting due to the CEP as of 2017. Following Koedel and Parsons (2021), the criteria we use to identify these states are (a) less than one percent of schools report an FRM share of 100 percent and (b) there is less than a five-percentage-point increase in the share of schools with missing FRM data from 2014 to 2017. The latter condition reflects the fact that in response to the CEP, some states have begun to report FRM data as missing. The 27 states that satisfy these criteria are: AL, AR, CA, CO, CT, FL, HI, IA, ID, IL, IN KS, KY, ME, MI, NC, NH, NJ, NY, OR, RI, TX, VA, VT, WA, WI, and WV.

For each of these states, we run regressions of the FM share on IPR (130), and the FRM share on IPR(185), as shown in Eqs. (6) and (7). Like in our analysis of the Missouri microdata, and under the maintained assumptions, coefficients of 1.0 on the IPR(X) variables would indicate that students' FM and FRM designations are aligned with the stated income requirements of the NSLP.

However, before turning to the 27-state expansion, we first establish comparability between our results using the Missouri administrative microdata (from above) and Missouri data taken from the CCD. In columns (1) and (2) of Table 5, we show results from regressions of the FM share on IPR(130), and the FRM share on IPR(185), respectively, using the CCD data to populate the FM and FRM variables for Missouri. Note that the Missouri data in the CCD are inclusive of CEP coding, so these results should correspond closely to the results in columns (6) and (8) of Table 2. Table 5 shows that this is indeed the case, confirming that the administrative NSLP data and CCD yield similar results in Missouri.

Columns (3) and (4) go on to show results from pooled regressions using the 27-state sample. We include state fixed effects in the pooled models to isolate within-state variation for identification, although as a practical matter this has no substantive bearing on the findings. The estimates using the 27-state sample are a close match to the Missouri estimates—the coefficient from the FM regression in particular is very similar, and the coefficient from the FRM regression in the larger sample is similar to, but somewhat lower than, the Missouri coefficient. This provides broad evidence that FM and FRM data overstate poverty rates under the condition that our assumptions on IPR(130) and IPR(185) also hold in the larger 27 state sample (more on this point below).

In addition, the pooled regressions in Table 5 obscure significant state-level heterogeneity in the estimated coefficients on IPR(130) and IPR(185). Fig. 2 illustrates this heterogeneity by plotting all 27 state coefficients and their error bands. For ease of presentation, states are ordered in each panel from the largest to smallest coefficient values. The range of estimates shown in Fig. 2 is striking. For example, in the FM regressions, the coefficient on IPR(130) ranges from a minimum of 0.50 (Arkansas), which implies FM is *underenrolled* by 50 percent, to a maximum of 1.75 (Rhode Island), which implies overenrollment by 75 percent. The range of coefficients in the FRM regressions is narrower, but still large, ranging from a minimum of 0.52 (Arkansas) to a maximum of 1.58 (Rhode Island).

This variability potentially reflects a number of factors that we

[19] The additional noise in the IPR-based regression is evidenced not only by the much larger standard error, but also the smaller R-squared value. An explanation is that the underlying IPR values are noisy estimates (based on data from just 25 households), and when they are shrunken by the Bayesian Kriging procedure, it restricts their variance substantially, reducing their explanatory power. The shrinkage is useful for our purposes because it enables us to recover unattenuated coefficients in our regressions, but the noise in the IPR estimates may limit their use in some applications. See discussion in Appendix A.

[20] For all 50 states and Washington DC, the CCD includes 99,165 schools in 2017. IPR estimates are available from the NCES for 99,156 of these schools—i.e., the coverage rate is essentially 100 percent.

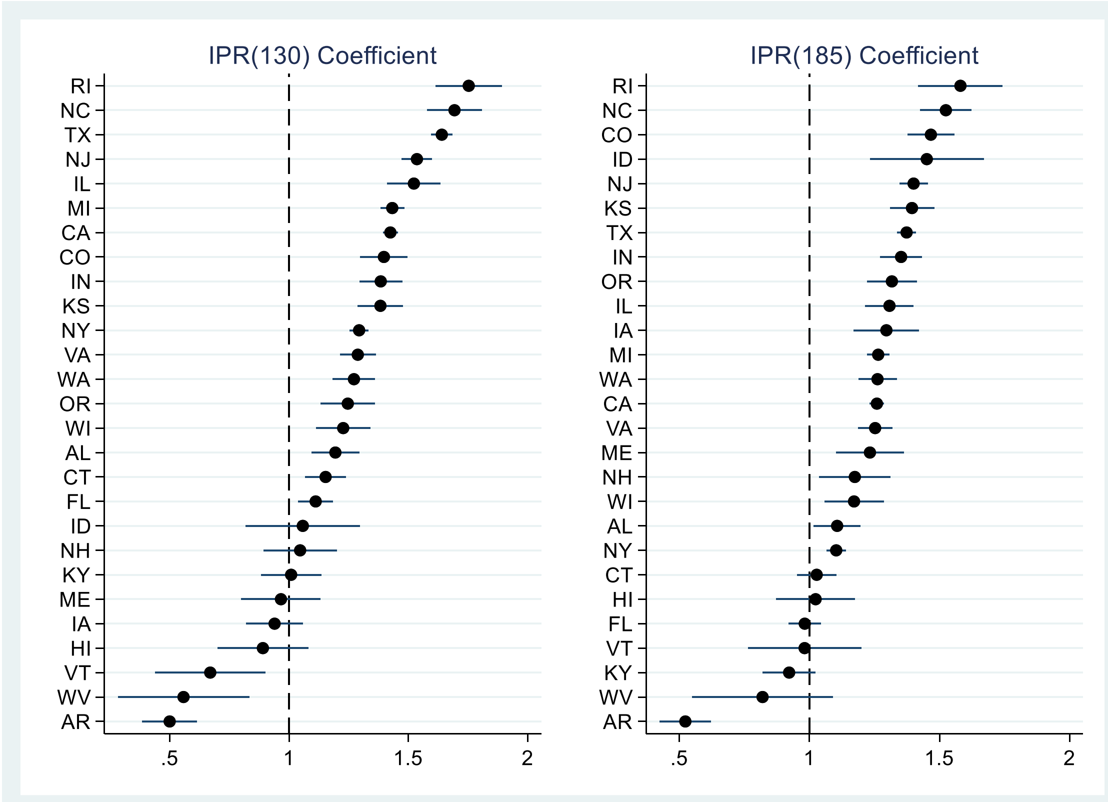

Fig. 2. Heterogeneity of the FM and FRM Regression Coefficients in the 27-State Sample.
Notes: The left panel shows the estimated coefficients from univariate regressions of the FM share on IPR(130) for the 27 states with non-CEP coded data, along with 95 percent confidence intervals. The right panel shows analogous coefficients and confidence intervals from univariate regressions of the FRM share on IPR(185). States are in descending order of the coefficient values in each panel. Data Source: Common Core of Data and SNP data, both from NCES, 2017.

cannot disentangle with our data. With regard to the FM and FRM enrollment data in the CCD, these include potential heterogeneity across states in enrollment processes and measurement error in the CCD itself. The former could reflect, for example, differences in leniency across states in districts' income-verification processes and/or differences in families' willingness to apply to the NSLP. The latter would include all reporting errors between the point of data collection in individual districts to the point of entry into the CCD. In addition, the degree to which there is variability in the accuracy of IPR(130) and IPR(185) across states is also uncertain. School attendance patterns with respect to geography could be different in other states. A related concern is that school locations within their communities relative to the local-area income distribution could differ systematically.[21] Regardless, the results in Fig. 2 raise concerns about using FM and FRM data from the CCD in multistate studies, which is a common practice. While we lack the data required to resolve the state-by-state discrepancies, future research to better understand the properties of the data and source(s) of heterogeneity across states would be valuable.

We also conduct an analog to the achievement-based analysis shown in Table 4 using the CCD data. While we do not have access to administrative data on student test scores in the multi-state sample, we can use data from the Stanford Education Data Archive (SEDA). SEDA contains district-level estimates of average standardized test scores in Math and English Language Arts for students in grades 3–8 throughout the U.S. The comparability across states is facilitated by linking the state tests and the National Assessment of Educational Progress (NAEP) to develop a common scale (Fahle et al., 2018; Fahle, Shear & Shores, 2019; Reardon, Kalogrides & Ho, 2021).

SEDA reports achievement at the district level, so we aggregate our poverty data accordingly and estimate regressions of district-level achievement on district-level measures of poverty using the 27-state sample.[22] We construct the district-level poverty shares as enrollment-weighted averages of the school-level poverty shares. We also add state fixed effects to our regressions, similarly to above, which yields the following analog to Eq. (8):

$$Y_{ks} = \varsigma_0 + P_{ks}\varsigma_1 + \nu_s + \eta_{ks} \tag{9}$$

In Eq. (9), Y_{ks} is the average math achievement level in district k in state s from SEDA, P_{ks} is the poverty measure of interest, and ν_s is a state fixed effect. In the Missouri-specific version of this model at the school level, we estimated it three times: defining P_{ks} as the DC share, IPR(130) estimate, and FM share. For the extended 27-state sample we do not observe the DC share, so we estimate the regression just twice—once

[21] As noted above, variation across states in the utilization of schools of choice, and private schools in particular, is an example of a specific factor that could affect the generalizability of our validity findings from Missouri. This is because the residences of private-school students are used to construct the SNP metrics (the households used by NCES include families with children but are not restricted to families with children in public schools), but these students will not affect poverty rates in public schools. All else equal, in states with larger private-school enrollment shares, there should be larger differences between IPR(130) and public-school poverty rates. Of note, Missouri's private-school enrollment share is above the national average, at 12.6 percent compared to 10.2 percent nationally as of 2017 (source: authors' calculations based on data as reported in de Brey et al., 2021).

[22] SEDA includes district-level average standardized math scores for a national sample of 9,728 out of 10,921 districts in the CCD (89 percent) in 2017. For our selected sample of 27 states, 6,221 out of the 6,853 districts in the CCD (91 percent) have math scores in SEDA in 2017.

Table 6
Relationships between student test scores and measured poverty, 27-state extended sample using the Common Core of Data (CCD) and the Stanford Education Data Archive (SEDA).

VARIABLES	(1) 2017 Dependent Variable: District Avg. Test Score	(2) 2017 Dependent Variable: District Avg. Test Score
District IPR(130)	−3.020***	
	(0.081)	
District FM Share		−1.404***
		(0.047)
Constant	0.815†††	0.364†††
	(0.043)	(0.036)
State Fixed Effects	Yes	Yes
Observations	6,221	6,221
R-squared	0.579	0.692

Notes: This is a national-level analog to Table 4 using data from CCD, SEDA, and SNP data from NCES. This table presents estimates from district-level univariate regressions where the dependent variable is the district average standardized test score, and the independent variables are IPR(130) and FM share in the district, respectively. The regressions are weighted by enrollment in each district. In each regression, we test the null hypothesis that the poverty-measure coefficient is zero; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For presentational consistency, we continue to denote statistical significance of the constant term at the 1, 5, and 10 percent levels using the same †††, ††, and † indicators from previous tables. Standard errors are in parentheses.
Data Source: Common Core of Data and SNP data, both from NCES, and SEDA data, 2017.

defining P_{ks} by IPR(130) and once defining it by the FM share. Following on our preceding analysis, all our SEDA-based regressions are weighted by student enrollment.

The results are shown in Table 6. As in our preceding analysis of the Missouri data, the coefficient from the regression using IPR(130) is much larger (i.e., more negative) than its analog using the FM share. Although both coefficients in Table 6 are larger than their comparison coefficients in the Missouri-specific analysis, the relative difference is similar.[23] This result further supports the conclusion that the FM share is not capturing the same level of poverty as IPR(130).

7. Implications for measuring poverty

Using a variety of data sources, samples, benchmarks, and outcomes, all of our findings consistently point to the conclusion that data from the NSLP do not measure student poverty accurately. Given this finding, it is natural to ask how states should respond. What should we do to measure poverty going forward?

Unfortunately, as is often the case, it is easier to identify limitations of existing approaches than new solutions (although we also note it has taken a long time, during a period of intensive policy use, for researchers to aggressively interrogate the properties of FRM data). And to be very clear, our findings do not lead to an obvious replacement for FRM data. Much more work is needed to understand the general problem of poverty measurement and vet alternative measures, including but not limited to the alternatives we consider here. Noting these qualifications, in the remainder of this section we review what we've learned about DC and SNP data in Missouri and how this can inform work to improve poverty measurement more broadly.

We view our administrative DC data as the most credible poverty data at our disposal. They are the least assumptive, there is reason to believe they are comprehensive for families with children (Schanzenbach, 2019), and DC policies in Missouri are relatively simple, making it more likely that eligibility criteria are adhered to. In fact, the origin story of this project begins with the fortuitious DC-data environment that makes Missouri well-suited to investigate the accuracy of students' FRM designations.

However, while we believe our results suggest promise with regard to the use of DC data to measure poverty, DC data are no panacea and may be less useful in other states. A significant concern with the broader application of DC data to measure poverty is that most states have a BBCE policy, and variability in these policies results in a range of poverty thresholds from 130 to 200 percent of the poverty line for directly certified students across states (United States Department of Agriculture, 2022). There are two implications of this. First, DC status conveys different information about the level of poverty in different states, which matters for both internal state policies and broader federal policies impacting multiple states. Second, it is unclear whether program participation and income enforcement in BBCE states are such that DC status will measure the income thresholds intended by state rules. For example, in some states participation in Medicaid can lead to direct certification (Blagg, Rainer & Waxman, 2019), but research shows that many Medicaid-eligible families do not participate (Sommers et al., 2012). Concerns have also been raised about the fidelity with which BBCE criteria are enforced (Blankley, 2019). Given the wide variability in states' BBCE policies—both in terms of their intended income-eligibility thresholds and the rules that govern direct certification—there is much to be learned from deeper investigations into the properties of DC data in other states.

In contrast to DC data, an appealing feature of IPR(X) (and related metrics based on the SNP data) is that it can be constructed to identify a common income level across all states. Our finding that IPR(130) is accurate, on average, when benchmarked against the DC data in Missouri is a promising data point regarding the value of this metric. However, our single study cannot provide sufficient evidence to advocate for its use more broadly. Future research using the SNP metrics—whether following our IPR(X) manipulation or otherwise—can help shed light on the measurement properties of the data. Our findings are at least suggestive that these data can be an informative piece of the puzzle, but more research is needed.

Ultimately, while our study offers some direction for future work, it

[23] The larger absolute values of the coefficients in Table 6 could be driven by a number of factors that differ in this portion of our analysis, including differences in state assessments that could differentially pick up differences in student poverty and disadvantage, factors related to SEDA's process for constructing comparable test scores across states, the impact of district aggregation on the estimates, and the related impact of variability in district size across states that leads to differential aggregation, among other possibilities. As this analysis is only supplementary to our main findings, and the absolute levels of the coefficients are not of first-order importance (it is their relative values that we care about), we did not thoroughly investigate the source(s) of the level differences in coefficient magnitudes.

is hard to describe the state of poverty measurement in public education as anything but bleak. FRM data are clearly problematic for measuring poverty—they are greatly oversubscribed, and pending more research, Fig. 2 suggests they may carry very different meanings across states. DC data seem promising in Missouri, but even if they are internally valid in Missouri and all other states (which is uncertain), they cannot be used consistently across states because direct certification identifies different income levels in different states. Moreover, DC policies (such as BBCE rules) are subject to change to achieve policy goals outside of the education system, and each time this happens, a DC-dependent data system will be disrupted, similarly to what happened when the NSLP introduced the CEP.

Regarding the SNP data, while we present some promising evidence on these data in Missouri, more research is needed to understand the generalizability of our findings. To the extent our findings generalize to other settings, SNP data could be useful to policymakers as a benchmark for understanding the information contained by other poverty measures (as we use the data here to understand FRM data) and to researchers who need a homogeneous measure of poverty across states. As we expand upon in Appendix A, the biggest concern with the IPR-based estimates of poverty is that they are inherently imprecise. The imprecision stems from the fact that their estimation is based on data from just 25 households per school. While studies such as ours can be designed to minimize the impacts of the imprecision, it may limit the use of IPR(130) and related matrics in some applications. For example, using IPR(130) to measure the poverty level for individual schools may be inadvisable even if it is an unbaised estimate on average, if the noise causes substantial prediction errors. And of course, SNP data are not suitable for use in any application that requires poverty information at the student level.

To end our discussion on a more positive note, these challenges present opportunities for researchers looking to improve education research and policy. FRM data are widely used in state funding and accountability policies to identify low-income children and widely used by researchers to measure poverty. Despite this, there has been relatively little work interrogating the properties of FRM data or developing alternative measures (where the former may explain the latter). Developing improved poverty measures will be challenging, and the data conditions are daunting, but efforts to improve poverty measurement could have large returns in the form of more accurate and impactful eduation research and policy.

8. Conclusion

We use detailed administrative data from Missouri to show that NSLP enrollment greatly overstates student poverty. For example, our estimates of the oversubscription rate for free meals in Missouri range from 35 to 50 percent. Under some additional assumptions, we extend our analysis to a larger sample of 27 states. This exercise suggests our findings are not unique to Missouri and potentially apply broadly across the U.S.

The excess enrollment in the NSLP is due in part to the introduction of community eligibility for free meals, which was rolled out nationally in the 2014–15 school year. However, we also show that there was excess enrollment in the NSLP prior to the CEP. We do not believe the historical inaccuracy of NSLP data is well understood, and this misunderstanding has implications for the development of new measures of poverty. For instance, some states and school districts are using a multiplier (above 1.0, with a commonly-advocated value of 1.6) to adjust DC-based poverty rates to match older FRM-based rates (e.g., see Croninger, Rice & Checovich, 2015; Grich, 2019). Although accuracy is not the only motivation for such adjustments (policy stability is another key consideration), policy documents suggest that it is not commonly understood that FRM-based poverty rates prior to the CEP are incorrect.[24]

Finally, to the extent that FRM-eligible students factor directly into states' funding and accountability policies, the process by which FRM data are generated, inclusive of potential heterogeneity across states and school districts, raises concerns about behavior that manipulates the underlying data (even if well-intentioned, such as educational administrators applying for the CEP or being more aggressive in eliciting and approving parental applications), which can affect the resulting fairness of these systems. With improved knowledge of the limitations of current and historical FRM data, we can make better choices in the development of new and more accurate measures of student poverty going forward.

Declaration of Competing Interest

We have no competing interests to declare.

Data availability

The authors do not have permission to share data.

Acknowledgement

We thank the Missouri Department of Elementary and Secondary Education for access to data, and Andrew Estep and Cheng Qian for research support. We gratefully acknowledge financial support from the Walton Family Foundation and 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 the funders, data providers, or institutions to which the author(s) are affiliated. All errors are our own.

Appendix A

Extended Analysis of IPR(130)

Over the course of showing the limitations of using FRM data to measure poverty, we provide the first external evidence of which we are aware on the accuracy of SNP-based poverty metrics, and IPR(130) in particular. IPR(130) is an appealing poverty measure for several reasons. Most notably, (a) it is conceptually well-grounded, (b) our empirical validation using the DC data suggests it is an accurate measure of poverty, at least on average and in Missouri, and (c) the underlying IPR data are published by NCES for virtually every school in the U.S., which means the data are widely available. As the limitations of FRM data become increasingly well-understood, researchers and policymakers will seek out alternative measures of poverty, and IPR(130) will be a prime candidate among sparse options. It is beyond the scope our study—both conceptually and in terms of feasibility due to data availability—to provide a detailed, national analysis of the prospects for using IPR(130) to measure poverty. However, in this appendix we

[24] The original source of the multiplier is the federal legislation that ushered in the CEP. The purpose is to approximate the share of FRM-eligible students in a CEP school based on the share who are directly certified. The multiplier is essentially what would be estimated from a regression very similar to the ones we run above of the FM share on DC share, but replacing the FM share with FRM share. The federal guidance establishing the multiplier explicitly references the link to the status quo of using FRM data to measure poverty: "using only the number of directly certified students would result in lower poverty percentages for Community Eligibility schools or LEAs" (U.S. Department of Education, 2014, p. 8). There is no mention that FRM data overstate poverty in federal documents or subsequent state documents that discuss similar multipliers (e.g., Croninger et al., 2015).

Appendix Table A1
Mean squared error analysis.

	2016	2017
FM share	0.051	0.054
IPR(130)	0.021	0.022
Observations	2,172	2,186

Notes: This table shows mean squared errors (MSEs) for the FM share and IPR (130). These MSE calculations assume the share of DC students reflects the true share of students at or below 130 percent of the poverty line. This is almost surely incorrect but is likely approximately accurate. Smaller values indicate less error.
Data Source: DESE administrative data and SNP data from NCES, 2016 and 2017.

provide some additional information and discussion based on our analysis in Missouri that we hope will improve understanding of the SNP data and motivate future research.

A.1. IPR(130) is a noisy measure of poverty

We begin by expanding on an issue we touch on in the main text, which is that IPR(130) is an inherently noisy measure of poverty for an individual school. This derives from the fact that the underlying IPR values from the NCES are estimated from data on just 25 households per school. The imprecision of the IPR(130) estimates is reflected in their relatively low variance. For example, Table 1 in the main text shows that while the DC share and IPR(130) have similar means in our dataset, the variance of IPR(130) across schools is less than 40 percent of the variance of the DC share (per the squared standard deviations: 0.0121 versus 0.0324). This, in turn, is because the IPR values are shrunken via the Bayesian Kriging procedure—the noise in the underlying estimates results in a strong pull toward the mean (prior), tightening the distribution of shrunken values.

Our evaluation framework is designed to leverage the shrinkage of the IPR values, and correspondingly IPR(130), into an analytic strength. Specifically, our use of a regression-based framework allows us to estimate average relationships between IPR(130) and the other poverty metrics—and test scores in some of our supplementary regressions—that do not suffer from attenuation bias (Chetty, Friedman & Rockoff, 2014; Jacob & Lefgren, 2008). However, the imprecision is still there and visible in two ways in our regression output from models involving IPR (130): our coefficients on IPR(130) have large standard errors and the R-squared values are relatively low. This can be seen throughout the tables in the main text.

While studies such as ours can be designed to minimize the impacts of the inherent imprecision of IPR(130), it may limit the use of IPR(130) in some applications. For example, using IPR(130) to measure the poverty level for individual schools may be inadvisable even if it is an unbaised estimate on average, if the noise causes substantial prediction errors.

Of course, such a statement is incomplete because it ignores the counterfactual. Ideally a better measure of poverty could be found, but what if data from the NSLP were the only alternative? We consider this question by comparing the predictive accuracy of IPR(130) to the FM share under the assumption that our DC data measure poverty accurately (this is an unverifiable assumption, but based on *a priori* expectations and the empirical analysis in our article, it is plausible). Under this assumption, we can compare the mean squared error (MSE) of IPR (130) and the FM share, relative to the DC share, across schools. The MSE is a particularly useful measure to analyze this question because it increases as the measures deviate from the DC share, whether due to bias *or* imprecision.

Table A1 shows that the MSE of the FM share is more than double the MSE of IPR(130). Even if the DC share is an imperfect measure of poverty itself, as long as it is at least roughly accurate, it seems difficult to argue based on these findings that IPR(130) is a worse measure of school poverty than the FM share. Of course, this brief analysis does not speak to whether researchers or policymakers should use IPR(130) to measure poverty in individual schools; but if the FM share is the only alternative, it does suggest IPR(130) is likely more accurate.

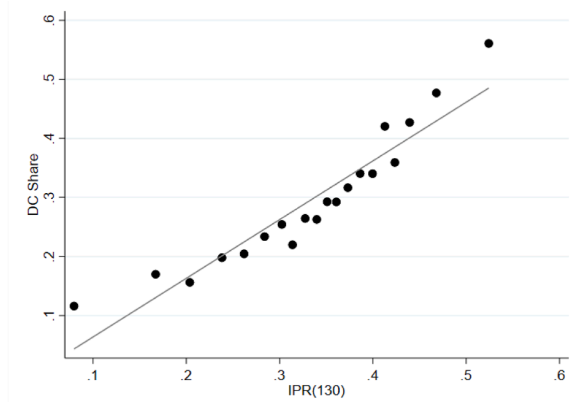

Fig. A1. IPR(130)-DC Share Binned Scatterplot for 2017.
Notes: We construct this chart by dividing IPR(130) into 20 equal-sized bins—each dot indicates the mean values of IPR(130) and the DC share within each bin. The full ranges of the IPR(130) and DC-share variables are 0–0.72 and 0.01–0.91, respectively. The regression line corresponding to column (2) of Table 2 is shown. Data Source: DESE administrative data and SNP data from NCES, 2017.

A.2. The disributional alignment of IPR(130) and the DC share is imperfect

Our primary validation regressions in Table 2 show that the DC share and IPR(130) convey very similar information, on average. But this does not imply full distributional alignment. Fig. A1 documents the relationship between IPR(130) and the DC share throughout the income distribution in Missouri. It shows a binned scatterplot of IPR(130) and the DC share using 2017 data, along with the corresponding regression line from column (2) of Table 2. Theoretically, if IPR(130) and the DC share are measuring identical constructs and contain no error, we would expect the data points to form a precise line. Instead, the data plot is slightly convex—the relationship between the two variables is flatter at lower poverty values and steeper at higher poverty values. We cannot identify the source(s) of the modest nonlinearity in the figure. One possibility is that there are subpopulations of students who are underrepresented relative to their poverty shares in programs that lead to direct certification and these students are clustered in the income distribution. This does not seem especially likely because the participation rate in SNAP among eligible families with children is estimated to be 100 percent (Schanzenbach, 2019), but some small deviations may exist. There may also be heterogeneity within the income distribution in the efficacy of the Kriging procedure used by NCES or in the efficacy of the procedure we use to recover IPR(130) from the underlying IPR values.

Our regression-based analysis in the main text is all conducted on average statewide, and thus it is sufficient for our purposes that IPR (130) and the DC share align on average (it also improves credibility that the nonlinearity documented in Fig. A1 is modest). That said, for other potential uses of SNP data (or DC data for that matter), the nonlinearity may be more concerning. For example, it can be a source of inaccuracy in income estimates for individual schools and raises concerns about the potential for systematic differences in the accuracy of IPR(130) for particular types of schools.

To consider this latter possibility in more detail, we explore the alignment between the DC share and IPR(130) using three splits of our sample, by: (1) schooling level (elementary/middle schools versus high schools), (2) urbanicity (rural versus urban/suburban schools), and (3)

Appendix Table A2
Univariate regressions of the DC share on IPR(130), school subgroups, unmatched samples.

	(1) Elem/Middle Schools	(2) High Schools	(3) Urban/Suburban Schools	(4) Rural Schools	(5) Charter Schools	(6) Traditional Schools
VARIABLES	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share
IPR(130)	1.118***	0.746***	1.141***	0.738***	1.298	0.968
	(0.036)	(0.063)	(0.051)	(0.042)	(0.289)	(0.034)
Constant	−0.046†††	−0.021	−0.072†††	0.040††	−0.062	−0.030†††
	(0.012)	(0.020)	(0.015)	(0.016)	(0.129)	(0.011)
Observations	1,612	574	877	1,309	64	2,122
R-squared	0.396	0.340	0.400	0.258	0.240	0.339

Notes: This table presents estimates from school level univariate regressions weighted by enrollment in 2017. In each regression, we test the null hypothesis that the poverty-measure coefficient is 1.0; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For completeness, we also report on the statistical significance of the constant term, where †††, ††, and † indicate the constant is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
Data Source: DESE administrative data and SNP data from NCES, 2017.

Appendix Table A3
Univariate regressions of the DC share on IPR(130), school subgroups, matched samples.

	(1)	(2)	(3)	(4)	(5)	(6)
	Matched Over Common Support Defined by High Schools		Matched Over Common Support Defined by Urban/ Suburban Schools		Matched Over Common Support Defined by Traditional Schools	
	Elem/Middle Schools	High Schools	Urban/Suburban Schools	Rural Schools	Charter Schools	Traditional Schools
VARIABLES	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share	2017 Dependent Variable: DC Share
IPR(130)	0.828***	0.746***	0.746***	0.940	1.298	1.256*
	(0.062)	(0.063)	(0.047)	(0.058)	(0.289)	(0.152)
Constant	−0.007	−0.021	0.018	−0.043††	−0.062	−0.020
	(0.018)	(0.020)	(0.013)	(0.020)	(0.129)	(0.055)
Observations	574	574	693	693	64	64
R-squared	0.297	0.340	0.276	0.313	0.240	0.434

Notes: This table presents estimates from school level univariate regressions weighted by enrollment in 2017. In each regression, we test the null hypothesis that the poverty-measure coefficient is 1.0; rejection of this null hypothesis at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively. For completeness, we also report on the statistical significance of the constant term, where †††, ††, and † indicate the constant is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
Data Source: DESE administrative data and SNP data from NCES, 2017.

school sector (charter versus traditional schools). Discrepancies between the DC share and IPR(130) within these subsamples could arise for a variety of reasons. One possibility is that the sizes of school catchment can differ along all three dimensions, which would affect the geospatial SNP metrics. Another is that for the charter schools, there could be a greater disconnect between geographic residence and school attendance. For the DC share, there could be students who are more or less likely to participate in social safety net programs in particular types of schools.

We report on the alignment between the DC share and IPR(130) for the data splits in Tables A2 and A3 (we show results for 2017 only for brevity; results for 2016 are similar). First, Table A2 shows results from simple regressions of the DC share on IPR(130) using each data subsample, which reveal coefficients that we can reject from 1.0 in four of six cases.[25] The misalignment is also not consistent across school groups within the data splits.

Table A3 investigates whether the misalignment is due to substantive reasons associated with the data splits, or alternatively, the bunching of particular types of schools within the income distribution combined with the nonlinearity illustrated in Fig. A1. To disentangle these mechanisms, we use a simple matching procedure to identify and compare groups of schools within each data split on a common support in the distribution of the DC share. We illustrate the procedure with the subsample comparison by schooling level. First, we run a simple, univariate probit predicting whether each school is a high school using the DC share. Next, we conduct a one-to-one match of high schools with elementary/middle schools using the predicted values (referred to as "propensity scores" in the matching literature). We impose a caliper of 0.01 and drop all schools without a match, which defines the common support. We then re-run the validation regressions using the matched sample of schools only. If the source of the subgroup misalignment is substantive, and not the nonlinearity, our findings will continue to differ between elementary/middle and high schools within the matched sample. Alternatively, if the findings are the same across school types in the matched sample, it would imply the nonlinearity—combined with the bunching of schools by type within the income distribution—is causing the misalignment in the unmatched data.

Table A3 shows that when we use the matched samples, our findings are generally similar across school groups within each split of the data. The coefficients still differ from 1.0 because in each setting we pull schools from only part of the income distribution via the matching procedure, and due to the nonlinearity, we should not expect a coefficient of 1.0 when we do this. However, the important comparisons are between the coefficients across school groups, within each data split, over the common support. Of the three data splits we consider, the coefficients within the urbanicity split differ the most from each other in

[25] The exceptions are for the charter/traditional schools split. For charter schools we fail to reject the null, but the test is not very informative due to the large standard error. For traditional schools, the overwhelming majority of our sample consists of traditional schools, so our estimate for this subgroup essentially replicates our estimate from Table 2 for the full sample.

Appendix Table A4
Mean squared error analysis, school subgroups.

	Elem/Middle Schools 2017	High Schools 2017	Urban/Suburban Schools 2017	Rural Schools 2017	Charter Schools 2017	Non Charter Schools 2017
FM share	0.049	0.068	0.058	0.052	0.132	0.052
IPR(130)	0.021	0.023	0.031	0.015	0.034	0.021
Observations	1,612	574	877	1,309	64	2,122

Notes: This table shows mean squared errors (MSEs) for the FM share and IPR(130) values for schools in 2017. These MSE calculations assume the share of DC students reflects the true fraction of students at or below 130 percent of the poverty line, which is almost surely incorrect but is likely approximately accurate. Smaller values indicate less error.
Data Source: DESE administrative data and SNP data from NCES, 2017.

Table A3, but even then, the matched-sample coefficients are much closer together than those from the unmatched sample in Table A2.[26] These findings indicate the primary source of misalignment in the subsamples is the nonlinearity, and not substantive factors that one might be concerned about that differ across school types.

Still, regardless of the cause, Tables A2 and A3 show that IPR(130) does not align with the DC share in the data subsamples. This highlights another concern with IPR(130), depending on the application, that merits attention in future research. It is also unclear if these results reflect a general property of the underlying IPR estimates or are idiosyncratic to Missouri.

Finally, in the same spirit as in Table A1, we can also add context to these findings by bringing in the FM share as a counterfactual measure. Table A4 shows MSE calculations for the data subsamples analogous to those in Table A1. The MSEs continue to show that IPR(130) matches the DC share much more closely than the FM share in the data subsamples, even though the correspondence between IPR(130) and the DC share is not as strong as in the full sample.

A.3. Summary

IPR(130) is a useful but imperfect measure of the poverty level of a school. Some of its limitations are inherent to its construction and likely independent of context, such as the fact that it is estimated using data from just 25 households. Other limitations—some of which we have touched on here, and others yet to be uncovered—may be context specific and/or idiosyncratic. In Missouri, we find that despite its limitations, IPR(130) is consistently a more accurate measure of school poverty than the FM share, highlighting the importance of considering the counterfactual in any assessment of its value. More broadly, if the goal is to produce a more accurate summative measure of poverty, overall or in subsamples of schools, using IPR(130) in conjunction with other measures—with an optimal weighting formula designed to minimize errors—may be promising. We conclude by noting these insights are based entirely on our analysis of Missouri data and could be refined considerably with more research on the properties of IPR(130) in other contexts.

Appendix B

Direct certification data in the CCD

The expansion of our analysis into other states in the main text would be more compelling if we could include comparisons of the IPR data to DC data, like in Missouri. Unfortunately, we lack administrative DC data from other states. We considered using DC data from the CCD for this expansion; however, we could not confirm their reliability enough to feel confident using them in our analysis. In this appendix, we document some of the concerns that came up with the DC data in the CCD for interested readers.

The biggest red flag in our investigation of the DC data in the CCD is that the data are not accurate for the one state we can credibly test: Missouri. To show this, we estimate a univariate regression analogous to those in the main text of schools' DC shares as reported in our administrative data on their DC shares as reported in the CCD. If the data elements in the Missouri microdata and the CCD are the same, we should anticipate a coefficient of 1.0 from this regression, but the coefficient is just 0.48. In contrast, when we run analogous comparative regressions using the FM and FRM data (i.e., regressing the shares from our administrative data on the shares in the CCD), we get coefficients very close to 1.0 (0.97 and 0.96, respectively), as anticipated. These results are reported in Appendix Table B1.[27]

We are not sure what is causing the discrepancy with the DC data in Table B1, but our Missouri microdata are surely more reliable because they are based on a direct merge of administrative files between

Appendix Table B1
DESE - CCD data relationship in MO, 2017.

VARIABLES	(1) Dependent variable: DC share from MO administrative data	(2) Dependent variable: FM share from MO administrative data	(3) Dependent variable: FRM share from MO administrative data
DC share from CCD	0.476***		
	(0.015)		
FM share from CCD		0.968***	
		(0.008)	
FRM share from CCD			0.964***
			(0.010)
Constant	0.146***	0.030***	0.030***
	(0.005)	(0.005)	(0.006)
Observations	2,183	2,183	2,183
R-squared	0.502	0.942	0.935

Notes: This table presents estimates from school level univariate regressions weighted by enrollment in 2017 for Missouri. *, **, and *** indicate the coefficient is statistically different from zero at the 1, 5, and 10 percent levels, respectively. Standard errors are in parentheses.
CCD=Common Core of Data.

[26] We also note that the urbanicity split has especially bad overlap because high-poverty schools are much more prevalent in the urban/suburban sample. This can be seen by the significant reduction in the total sample size in the matched data in Table A3 relative to Table A2—many schools are dropped even in the "treatment pool" of urban/suburban schools because there are no rural school matches.

[27] These findings for the FM and FRM data are consistent with the fact that in the main text, we obtain similar results from regressions of the FM and FRM shares on IPR(130) and IPR(185), respectively, regardless of whether we use our administrative data or data from the CCD (see Tables 2 and 5).

Appendix Fig. B1. Availability of DC and FRM Data in the CCD, 2017.
Note: States are arranged in descending order by the percent of schools with DC information provided.

agencies. It may be that DC data collection and reporting procedures in the CCD are newer than for FRM data and not as carefully vetted, or there may be some other explanation for the discrepancy. But whatever the cause, the lack of alignment in Missouri is cause for concern about the DC data in the CCD more broadly. To be clear, we cannot identify concrete discrepancies in any other states' DC data in the CCD because we do not have access to administrative data for comparison, but we also cannot confirm the reliability of their DC data.

We also document DC data availability more broadly in the CCD, focusing on the 2017 data file (which is the file relevant to our analysis). For the majority of states, they report either no DC information, or very little DC information. This is illustrated in Appendix Fig. B1, which shows the percent of schools in each state that report DC and FRM information in 2017. Only 26 states had at least one school reporting the number of DC students in the CCD, and just 17 states had this information for more than 95 percent of schools. In comparison, 47 states had at least one school reporting the number of FRM students, and 42 of these states had this information for more than 95 percent of schools. Although the lack of coverage of DC data in the CCD does not directly indicate problems with the data that are provided, it does raise concerns that merit testing.

As a final point of information, we pulled analogous DC data from the most recent CCD release (for 2021) to see if data conditions are improving. By 2021, more states report at least some DC data (30 instead of 26), and more states have at least 95 percent coverage (21 instead of 17). This suggests the DC data are improving in the CCD, albeit slowly, and with the caveat that we can only see an expansion of data availability and cannot test its reliability (this is true even in Missouri, where we do not have access to administrative DC data in more recent years).[28]

Ultimately, between the issues we found with the Missouri data and the incomplete DC data coverage in the CCD more broadly, we did not feel confident carrying our analysis forward in other states based on DC data from the CCD.

[28] At the same time, FRM data are becoming less available (the number of states reporting at least some FRM data in 2021 declined from 47 to 42, and the number of states with at least 95 percent coverage fell from 42 to 37).

References

Bass, D. N. (2010). Fraud in the lunchroom? Federal school-lunch program may not be a reliable measure of poverty. *Education Next, 10*(1), 67–72.

Blagg, K., Rainer, M., & Waxman, E. (2019). *How restricting categorical eligibility for SNAP affects access to free school meals. Policy report*. Washington, DC: Urban Institute.

Blankley, B. (2019). *Public comments close for new SNAP benefits rule closing eligibility loophole*. The Center Square (09.24.2019). Retreived on 01.31.2022 from https://www.thecentersquare.com/national/public-comments-close-for-new-snap-benefits-rule-closing-eligibility-loophole/article_37be5546-decc-11e9-82ec-87edc7c4bef6.html.

Burghardt, J., Silva, T., & Hulsey, L. (2004). *Case Study of National School Lunch Program Verification Outcomes in Large Metropolitan School Districts. Nutrition Assistance Program Report Series*. US Department of Agriculture. Report No. CN-04-AV3.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–2679.

Chingos, M. M. (2016). No more free lunch for education policymakers and researchers. *Evidence Speaks Reports, 1*(20), 1–4.

Cressie, N. (1989). Geostatistics. *The American Statistician, 43*(4), 197–202.

Cressie, N. (1993). *Statistics for spatial data*. New Jersey, NJ: John Wiley & Sons.

Croninger, R. G., Rice, J. K., & Checovich, L. (2015). *Evaluation of the use of free and reduced-price meal eligibility as a proxy for identifying economically disadvantaged students. Alternative measures and recommendations*. Denver, CO: Augenblick, Palaich & Associates.

Davis, W., & Musaddiq, T. (2019). Estimating the effects of universal free school meal enrollment on child health: Evidence from the community eligibility provision in georgia schools. Working paper.

de Brey, C., Snyder, T. D., Zhang, A., & Dillow, S. A. (2021). *Digest of education statistics 2019*. National Center for Education Statistics. NCES 2021-009.

Domina, T., Pharris-Ciurej, N., Penner, A. M., Penner, E. 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.

Fahle, E. M., Shear, B. R., & Shores, K. A. (2019). Assessment for monitoring of education systems: The US example. *The ANNALS of the American Academy of Political and Social Science, 683*(1), 58–74.

Fahle, E.M., Shear, B.R., Kalogrides, D., Reardon, S.F., DiSalvo, R., & Ho, A.D. (2018). Stanford education data archive: Technical documentation (version 2.1). Retrieved from http://purl.stanford.edu/db586ns4974.

Geverdt, D. (2019). *Education demographic and geographic estimates program (EDGE): School neighborhood poverty estimates, 2016-2017 (NCES 2019-112)*. Washington, DC: National Center for Education Statistics. U.S. Department of EducationRetrieved from https://nces.ed.gov/pubsearch/.

Geverdt, D., & Nixon, L. (2018). *Sidestepping the box: Designing a supplemental poverty indicator for school neighborhoods (NCES 2017-039)*. Washington, DC: National Center for Education Statistics. U.S. Department of EducationRetrieved from http://nces.ed.gov/pubsearch/.

Gindling, T.H., Mata, C., Kitchin, J., & Avila, E. (2018). Some causes of undercount of low income students under the Community Eligibility Provision in Baltimore City Public Schools (No. 18-01). Department of Economics Working Paper Series Paper.

Gordon, N., & Ruffini, K. (2021). Schoolwide free meals and student discipline: Effects of the community eligibility provision. *Education Finance and Policy, 16*(3), 418–442.

Greenberg, E. (2018). *New measures of student poverty: Replacing free and reduced-price lunch status based on household forms with direct certification*. Urban Institute.

Greenberg, E., Blagg, K., & Rainer, M. (2019). *Measuring student poverty: Developing accurate counts for school funding, accountability, and research*. Urban Institute. Research Report.

Grich, R. (2019). *New strategies for measuring poverty in schools*. Explainer. FutureEd. Retrieved from: https://www.future-ed.org/how-states-measure-poverty-in-schools/.

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

Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics, 26*(1), 101–136.

Koedel, C., & Parsons, E. (2021). The effect of the Community Eligibility Provision on the ability of free and reduced-price meal data to identify disadvantaged students. *Educational Evaluation and Policy Analysis, 43*(1), 3–31.

Lichter, D., Sanders, S., & Johnson, K. (2015). *Behind at the starting line: Poverty among Hispanic infants*. Durham, NH: University of New Hampshire, Carsey School of Public Policy. Retrieved from https://scholars.unh.edu/cgi/viewcontent.cgi?article=1250&context=carsey.

Lundborg, P., Rooth, D., & Alex-Petersen, J. (2022). Long-term effects of childhood nutrition: Evidence from a school lunch reform. *The Review of Economic Studies, 89*(2), 876–908.

Massachusetts Department of Elementary and Secondary Education. (2017). *Policy Report from the Massachusetts Department of Elementary and Secondary Education*.

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. (2019). Accounting for student disadvantage in value-added models. *Journal of Educational and Behavioral Statistics, 44*(2), 144–179.

Reardon, S. F., Kalogrides, D., & Ho, A. D. (2021). Validation methods for aggregate-level test scale linking: A case study mapping school district test score distributions to a common scale. *Journal of Educational and Behavioral Statistics, 46*(2), 138–167.

Ruffini, K. (2021). Universal access to free school meals and student achievement: Evidence from the community eligibility provision. *Journal of Human Resources*. https://doi.org/10.3368/jhr.57.3.0518-9509R3

Sandstrom, H., Huerta, S., & Loprest, P. (2014). *Understanding the dynamics of disconnection from employment and assistance: Final report*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B, 13*(2), 238–241.

Schanzenbach, D. W. (2019). Exploring options to improve the Supplemental Nutrition Assistance Program (SNAP). *The ANNALS of the American Academy of Political and Social Science, 686*(1), 204–228.

Schwartz, A. E., & Rothbart, M. W. (2020). Let them eat lunch: The impact of universal free meals on student performance. *Journal of Policy Analysis and Management, 39*(2), 376–410.

Sommers, B. D., Tomasi, M. R., Swartz, K., & Epstein, A. M. (2012). Reasons for the wide variation in Medicaid participation rates among states hold lessons for coverage expansion in 2014. *Health Affairs, 31*(5), 909–919.

U.S. Department of Education. (2014). *Guidance: The community eligibility provision and selected requirements under title I, Part A of the Elementary and Secondary Education Act of 1965, as amended*. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education.

Williams, S. (2013). Public assistance participation among U.S. children in poverty, 2010 (FP-13-02). Working paper, National Center for Family & Marriage.

United States Deparment of Agriculture. (2022). *Broad-based categorical eligibility (Last update: January 2022). Info sheet*. Washington DC: United States Department of Agriculture. Retrieved 02.01.2022 at: https://www.fns.usda.gov/snap/broad-based-categorical-eligibility.

Zedlewski, S. R., & Martinez-Schiferl, M. (2010). Low-income Hispanic children need both private and public food assistance. *Policy brief*. Urban Institute.