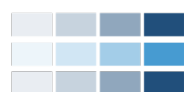


ESSER and Student Achievement: Assessing the Impacts of the Largest One-Time Federal Investment in K12 Schools

Dan Goldhaber
Grace Falken

March 2025

WORKING PAPER No. 301-0325-2



CALDER

National Center for Analysis of
Longitudinal Data in Education Research



ESSER and Student Achievement: Assessing the Impacts of the Largest One-Time Federal Investment in K12 Schools

Dan Goldhaber

*American Institutes for Research / CALDER
University of Washington*

Grace Falken

University of Washington

Contents

Contents.....	i
Acknowledgments	ii
Abstract	iii
1. Introduction	1
2. The Allocation of ESSER Funding and Our Identification Strategy.....	4
3. Data and Empirical Specifications	9
4. Results	17
5. Placebo and Robustness Tests.....	19
6. Heterogeneity	22
7. Discussion & Conclusion	23
References	28
Tables and Figures.....	35
Appendix: Supplemental Results, Figures, and Tables	45

Acknowledgments

We thank Luke Caldwell for helping gather data for some analyses presented here and Michael DeArmond for editorial suggestions. Marguerite Roza, Maggie Cicco and the Edunomics team provided data on ESSER allocations that are a key part of our analysis. We benefited from comments by Ben Backes, Nora Gordon, Eric Hanushek, Thomas Kane, Kirabo Jackson, Josh McGee, Emily Morton, and Sean Reardon. We are especially grateful for advice we received from James Cowan at multiple points in the development of the paper. This paper was supported by the Center for Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations (see <https://caldercenter.org/about-calder>).

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

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

ESSER and Student Achievement: Assessing the Impacts of the Largest One-Time Federal Investment in K12 Schools

Dan Goldhaber & Grace Falken

CALDER Working Paper No. 301-0325-2

March 2025

Abstract

We estimate the effects of federal pandemic-relief funding (ESSER) for K12 district-level student achievement in 2023. We rely on test data from almost 5,000 school districts across 28 states. Our identification strategy exploits variation in 2022-23 ESSER spending attributable to its allocation rules. We find that a \$1,000 increase in ESSER spending per pupil led to a statistically significant increase in math test performance of 0.007 standard deviations. The impact of ESSER on ELA was smaller, .002 standard deviations, and not statistically significant. ESSER differentially impacted districts across poverty levels and urbanicity.

1. Introduction

During the COVID-19 pandemic, the federal government invested an unprecedented amount of funding in K12 schools through the Elementary and Secondary School Emergency Relief Fund (ESSER).¹ Totalling nearly \$200 billion, ESSER represents the largest one-time investment in K12 schools in American history. Ninety percent of ESSER went directly to local school districts to support the reopening of in-person learning and academic recovery. While the funding was largely unrestricted, 20 percent of the third wave of ESSER was designated to address pandemic-induced “learning loss” (Office of Elementary & Secondary Education, 2021). Despite this substantial investment, there is currently limited evidence about whether ESSER funds helped students catch up from the large declines in test achievement they experienced during the pandemic (National Center for Education Statistics, 2022a, 2022b).

Whether ESSER bolstered student achievement is a significant policy question, particularly given the 2024 release of the National Assessment for Educational Progress (NAEP) showing that achievement in grades 4 and 8 in both reading and math remains far below pre-pandemic levels; in reading, the NAEP results show achievement has continued to erode since the pandemic (National Center for Education Statistics, 2025). This lackluster recovery aligns with earlier polls of 2024 voters in four battleground states who were skeptical of the impact of ESSER funding (Stanford, 2023). More broadly, whether ESSER impacted student test results

¹ There were three waves of ESSER funding: ESSER I included \$13.2 billion and was approved in March 2020 as part of the Education Stabilization Fund through the Coronavirus Aid Relief and Economic Security (CARES) Act; ESSER II included \$54.3 billion and was approved in December 2020 as part of the Coronavirus Response and Relief Supplemental Appropriations Act (CRRSA); and ESSER III included \$122 billion was approved in March 2021 as part of the American Rescue Plan (ARP) Act (Office of Elementary & Secondary Education, 2024).

speaks to long-standing debates about the extent to which increased school funding leads to improved student achievement (Hanushek, 1989, 1994; Hedges et al., 1994).²

In this paper, we answer the question of whether ESSER funding had a measurable impact on student test achievement in the 2022-23 school year. We use district-level data from 28 states and our analytic framework yields plausibly causal estimates. In particular, to address concerns that spending may be endogenous, we instrument for ESSER spending using the share of children in a district's geographic area counted as formula-eligible for Title I funding. Although Title I allocations depend on other features of districts and states (Gordon & Reber, 2023), the main determinants are the formula-eligible count (FEC) of children in the district and the formula-eligible percent (FEP) of the local, school-aged population.³ The identifying assumption of our design is that, after we control for district resources, poverty levels, and demographic characteristics in 2022-23, differences in funding attributable to FEP are exogenous to the characteristics of the school district and its students. We leverage variation in funding due to the reporting delay in FEP,⁴ differences between community demographics and those of students enrolled in public schools, and measurement error in FEC and FEP to isolate the impacts of ESSER on student achievement.

We find consistent, significant effects of ESSER spending on student achievement in math; the results in English language arts (ELA), however, are more sensitive to model specification. Our preferred models, which include state fixed effects, suggest that a \$1,000

² The most recent research on the question “does money matter?” finds increased spending improves outcomes such as test scores, educational attainment, wages during adulthood, and reduces the likelihood of being arrested (Baron et al., 2024; Jackson et al., 2016, 2021; Kreisman & Steinberg, 2019; Lafortune et al., 2018).

³ For clarity, and following the convention of Gordon and Reber (2023) we refer to the percent of children who are formula-eligible in district as the formula-eligible percent (FEP) and the number of formula-eligible children as the formula-eligible count (FEC).

⁴ The data on FEC and FEP are reported on a lag such that data from 2018 determined Title I funding for the 2020-21 school year; this delay translates to variation in funding that is explained by changes in poverty over time and not reflective of current-year poverty conditions.

increase in ESSER spending per pupil resulted in a statistically significant increase of about .007 standard deviations in math and a statistically insignificant .002 increase in ELA achievement.⁵ The math estimate is broadly consistent with evidence from a meta-analysis by Jackson and Mackevicius (2024) who find that the average effect of a \$1,000 increase in per-pupil spending over four years is about .032 standard deviations, or approximately 0.008 in a single year.⁶ A placebo test designed to assess omitted variables bias and using alternative measures of student poverty support evidence of ESSER effects in math.

We also find some evidence of heterogeneity in ESSER spending effects. Specifically, our results suggest that the impacts of ESSER were slightly larger for higher-poverty districts and smaller in suburbs relative to other urbanities. We do not observe differences according to student demographics or pre-pandemic district revenue. However, these conclusions are tenuous given our limited power to detect effects.

Our work extends the prior causal research on the impacts of funding in several ways. First, prior work relies on identifying variation in spending arising from school equity and adequacy lawsuits (e.g., Jackson et al. 2016) or variation within single states (e.g., Kreisman and Steinberg 2019). We instead rely on exogenous variation linked to Title I funding, which is distinct both in centering federal revenue variation instead of state or local revenue and in allowing us to assess the impact of spending across and within states and across a much broader set of school districts.⁷ Second, because ESSER funds are proportionate to Title I, we focus on

⁵ As we discuss below, the math results are robust to using several alternative controls for contemporaneous poverty, and, in both subjects, a placebo test of applying our specification to 2018-19 data, in which we would expect to capture only the impact of Title I funding, raises no cause for concern.

⁶ Jackson and Mackevicius document that these effects are larger for higher poverty student populations, but they find little evidence of heterogeneity across capital and operating spending, for different spending levels, or by geography. See also Handel and Hanushek's (2023) meta-analysis, which reaches similar conclusions about the overall impact of increases in school spending.

⁷ Cascio et al. (2013) examine the effect of the introduction of Title I in Southern states on future spending and student dropout rates. Matsudaira et al. (2012) look at the effects of variation in Title I funds received at the school-

the neediest school districts (Fahle et al., 2023; Kuhfeld et al., 2022), a distinct margin of analysis from past research. Finally, to our knowledge, this is one of only two studies that focus on a time-limited funding increase (districts were required to obligate the last wave of ESSER funding by September 2024).⁸ Our findings are broadly consistent with similar studies by Dewey, et al. (2024, 2025), which we discuss in greater detail below.

2. The Allocation of ESSER Funding and Our Identification Strategy

Central to our study is the fact that the federal government allocated ESSER funding in proportion to Title I, providing more support for higher-need schools and districts. As a result, while Detroit Public School District received about \$25,800 per pupil across all waves of ESSER, the Grosse Pointe district (a nearby suburb) only received about \$860 per pupil. These two districts, and, more generally, districts with different ESSER allocations or spending rates may also have distinct underlying needs related to both funding and achievement, posing a challenge for assessing the causal impact of ESSER spending on test achievement. For example, higher-poverty districts like Detroit tend to be in communities that had higher rates of COVID infection and deaths (Chen & Krieger, 2021; Finch & Hernández Finch, 2020; Karmakar et al., 2021), more negative impacts on mental health and well-being (Hall et al., 2022), higher unemployment (Tang et al., 2022), greater likelihood of child maltreatment (Wolf et al., 2024), and more remote schooling (Goldhaber et al., 2023). These and other factors likely exacerbated learning loss during the pandemic and dampened academic recovery. If these types of factors influence ESSER spending and are not accounted for in achievement models, naïve ordinary

level in a large, urban district using a regression discontinuity. Little other research appears to explore the impact of federal funding variation on student achievement deploying causal methods, but Goldhaber et al. (2025) exploit the same source of variation that we do here in assessing how ESSER funding affects school district staffing decisions.

⁸ As we describe in more detail below, there is relatively little systematic evidence about how school districts used ESSER funding outside of broad expenditure categories. Moreover, official reports provide limited evidence on the resources that were purchased *because of* the additional ESSER funding given the fungibility of school spending (Gordon, 2004).

least squares (OLS) estimates of the effects of ESSER will be biased downward.⁹ We could also imagine scenarios leading to an upward bias of an OLS estimate. For example, districts that spend ESSER faster could also have greater capacity to implement academic recovery interventions. It is also possible that states or communities might have invested in academic recovery in ways that are not accounted for by the controls in the model but are correlated with ESSER spending.¹⁰

One solution to these challenges is to use ESSER *allocations* in place of spending, but this too is potentially problematic given the hold harmless rules that influence allocations. When districts experience declines in FEC that drop their Title I funding beyond certain levels, their allocation is held harmless at between 85 and 95 percent of their previous year Title I allocation (Gordon & Reber, 2023). The likelihood of being held harmless could be correlated with achievement changes not well-captured by factors accounted for in statistical models.¹¹ Given this, we address concerns about potential bias from using ESSER spending by exploiting variation in ESSER spending attributable to its *allocation rules* and their relationship to Title I.¹² Title I combines four grants, each with distinct rules determining eligibility and formula adjustments. The core determinants of each Title I grant, however, are the number of FEC and FEP in a district's geographic area. Given the uneven distribution of poverty, there is

⁹ For instance, there is evidence that more time spent in remote schooling is associated with the chronic absenteeism that is making academic recovery challenging (Dee, 2024; Goldhaber et al., 2023).

¹⁰ As an example, if states or communities invested in mental health initiatives for students that targeted low-income areas and are funded outside of district budgets—one report estimates that as much as \$1 billion in state-reserved funds were planned to target mental health and well-being supports (Council of Chief State School Officers, 2022, 2024)—the impacts of that programming would be correlated with ESSER and also impact achievement.

¹¹ For instance, as a district becomes more affluent, low-income families could be priced out of the district, triggering hold harmless provisions, meaning Title I allocations do not decline proportionately to the drop in FEP. If this increasing affluence is correlated with student achievement, then OLS estimates would be upwardly biased.

¹² It is worth noting that, in some states, grant allocations are not 100% proportional due to the distribution of state-reserved ESSER funds to districts. In Washington, for example, the state education agency granted each district that did not receive Title I funding in the pandemic years (and thus would not have otherwise received ESSER funds) with \$75,000 of ESSER funds from the state withholding.

considerable variation in how much ESSER funding districts receive. For example, the mean ESSER allocation per pupil in high-poverty states like Mississippi, about \$6,700, is almost four times greater than the mean in the low-poverty state of Connecticut, which received about \$1,800 per pupil.

Researchers have tried to assess the impacts of Title I on students since the inception of the program. Borman and D'Agostino (1996) summarize the early literature on this question in a meta-analysis, finding modest, positive effects of Title I funding that grew as the program became more established. Several studies have used more rigorous research methods including regression discontinuity designs to identify impacts of qualifying for Title I within individual districts, finding no positive impacts on student achievement (Matsudaira et al., 2012; van der Klaauw, 2008; Weinstein et al., 2009). Cascio et al. (2013) leverages an event-study design and a 2SLS model that instruments for the change in federal revenue at the inception of Title I with poverty changes over time, an approach that is similar to that of this paper. Rather than look at test achievement, Cascio et al (2013) focus on the impacts of the new funding on district finance and high school dropout rates; they find that Title I improved (reduced) dropout rates for White students but not Black students. Johnson (2015), using longitudinal data with birth-cohort and district fixed effects, finds positive impacts of Title I on educational attainment and outcomes in adulthood, but also does not explore impacts on test achievement.

Our paper is most closely related to Dewey et al. (2024) and its follow-up (Dewey et al., 2025). Dewey et al. (2024) base their analysis on a similar sample of districts and use the same achievement data we use (described in more detail below). Their ESSER measures of spending are from Burbio, a private service that collected proposed ESSER spending and updated the data as realized spending data became available. It is unclear the extent to which the Burbio data was

updated by time of analysis in Dewey et al. (2023, 2024) and would thus be similar to or different from the spending data we utilized, which we collected from the 28 states in our sample.¹³ Overall, our findings fall in a similar range: Dewey et al. (2024) find a \$1,000 increase in ESSER spending significantly improved math test scores by .009 SDs and ELA by .005 SDs.

Dewey et al. (2025) is slightly broader in scope than our study, examining changes in achievement between 2022 and 2024 in 43 states. The major differences from the current project are twofold: (1) Dewey et al. (2025) consider two years of change in achievement for a larger sample, and (2) they use ESSER *allocations* as the predictor of interest, which likely underestimates the full effect of ESSER since districts still had more time to spend ESSER funds following spring test administration. Dewey et al. (2025) find that ESSER allocations improved math test gains between 2022 and 2024 by .005 standard deviations (SDs) and ELA gains by .006 SDs, suggesting more modest single-year gains of .003 in each subject.

Our identification strategy hinges on the measures determining Title I. To help clarify our approach, we present a timeline of the relevant measures and events, associated data for FEC and FEP, and impacted ESSER allocations in Figure 1. ESSER I allocations in 2020 were determined by Title I allocations for the 2019-20 school year, which themselves were anchored to FEC and FEP data from 2017. Similarly, ESSER II and III allocations in 2020 and 2021, respectively, were determined by Title I for 2020-21, which was anchored to FEC and FEP data from 2018. Because districts had spent most of ESSER I and more than half of ESSER II by the start of the 2022-23 school year, our analysis focuses on the effects of ESSER II and III.

A key aspect of the Title I allocation formula is that, as a result of several allocation rules, there can be meaningful differences in Title I allocations for districts with comparable FEC

¹³ Because of these data differences, their sample includes 30 states in the Stanford Education Data Archive (discussed below) while ours includes 28. We are missing data from Alabama and Oklahoma.

and FEP (Gordon & Reber, 2023). First, state minimums (a guaranteed funding floor) inflate the Title I dollars allocated per FEC in less-populated states relative to more densely populated states. Second, dollars allocated per FEC are adjusted based on state per-pupil expenditure (SPPE), so states that spend less money per student have, all else equal, lower allocations per FEC. Third, Education Finance Incentive Grants (EFIG)—one of the four grants under Title I—adjust state allocations according to SPPE relative to state per capita income and also how equally funds are distributed across districts, so allocations per FEC differ across grants. Finally, Title I allocations also vary per FEC within states due to (1) differences in the weighting of FEC for allocations above certain FEP thresholds and (2) adjustments to allocations to maintain hold-harmless provisions.¹⁴

As we describe in more detail below, the key mechanism we rely on to estimate the effects of ESSER on student test achievement is the fact that ESSER allocations are determined by *prior* measures of district poverty (based on the FEC and FEP data year). Our central assumption is that *conditional on current district characteristics*—such as demographics and the share of free/reduced price lunch (FRPL) eligible students in the 2022-23 school year—the extra funding districts receive because of the above allocation rules are uncorrelated with unobserved district-level factors that influence student achievement.¹⁵

¹⁴ Gordon and Reber (2023) delve into these factors in greater detail, but generally allocations are determined by an iterative process where districts are guaranteed to maintain some level of their prior-year funding if their initial allocation declines (due, for example, to falling below the qualification threshold of a grant). Because districts then receive a “boost” from their initial allocation to remain at a certain level, and the overall budget constraint of the program, districts unaffected by hold harmless clauses must make up the budget gap by receiving funds slightly lower than their initial allocation.

¹⁵ We use data on FRPL eligibility from the Common Core of Data which is noted as a flawed source in part because if districts participate in the Community Eligibility Program (providing free lunch to all students regardless of eligibility) then participation counts overstate eligibility in that district (National Center for Education Statistics, 2020). Because FRPL is thus top-constrained in some districts we additionally re-estimate our models dropping those censored districts—a robustness check we describe in greater detail below.

We provide an illustration of this source of cross-sectional identifying variation in Figure 2, in which we plot ESSER II and III allocations per pupil against the district FEP in Panel A and the relationship between district FRPL eligible students and FEP in Panel B. Panel A illustrates that while there is a positive correlation between FEP and ESSER (.56 for the school districts in the figure and .62 in our sample, described below), there is also variation in spending for districts with the same FEP. Panel B, which plots 2018 FEP (which determined Title I in 2020-21) against the percentage of FRPL eligible students in each district in 2022-23, shows a much noisier connection between FRPL and FEP. The time lag in FEP leads to a mismatch with current-year poverty levels, but this noise also comes from differences in the demographics of children in a district area relative to children enrolled in the district. Additionally, mismatches between these two measures of poverty may arise from the fact that while FRPL numbers are a reflection of the *observed* share of enrolled students in the district, FEP and FEC are *estimates* calculated based on several data sources and thus inherently have some amount of error (U.S. Census Bureau, 2023).

The central argument that we make is that while the Title I and ESSER spending (and allocations received by districts) may be correlated to underlying needs (and cost of educating students in each district), the variation in ESSER spending associated with the time lag and changed district demographics should not be correlated to these underlying needs. This variation therefore allows us to assess the benefits of additional funding received that is *not* directly connected with the true cost of education in those districts. As we describe below, we test this assumption by estimating a placebo test on pre-pandemic achievement.

3. Data and Empirical Specifications

3.1 Data sources and measures

Our analyses rely on five main sources of publicly available data: (1) district characteristics such as demographics, staffing, finance, and enrollment from the Common Core of Data (CCD); (2) district-level FEC and FEP in 2018 used to determine Title I allocations for the 2020-21 school year and for constructing our instrument;¹⁶ (3) data from the Stanford Education Data Archive 2023 (SEDA) to measure student achievement, (4) ESSER allocations from the ESSER Expenditure Dashboard, maintained by the Edunomics Lab at Georgetown University, with some exceptions we describe below; and (5) ESSER spending data for the 2022-23 school year which we individually requested and collected from 28 states.¹⁷ Note that the SEDA data exclude several large, populous states including Texas and New York, so the findings we describe may not generalize to the U.S. as a whole. We provide more details about SEDA data and the other datasets below and list the included states by subject in Appendix Table A.1.

For general information about districts, we use files from the CCD (Common Core of Data 2023). Specifically, we use CCD data to observe urbanicity, enrollment, and student demographics, all for the 2022-23 school year, and total revenue per pupil for the 2019-20 school year. Because the CCD maintains records for all districts in the country, these data allow us to describe how our analytic sample of districts compares to all other districts in the U.S. and consider heterogeneity within districts we observe. Most important to our analysis are the CCD data on FRPL qualification rates in a district. These data are not completely accurate for high-poverty school districts qualifying for the Community Eligibility Provision, as districts are

¹⁶ These data are available upon request from the Department of Education Office of Formula Grants/School Support and Accountability.

¹⁷ We also use data on instruction modality at the district-by-week level from the Return to Learn Tracker maintained by the American Enterprise Institute (2021). The sample for this dataset is limited to districts with three or more schools.

allowed to provide all students with free lunches and thus report 100% FRPL, even when the actual percentage is lower (National Center for Education Statistics, 2020). Due to this constraint, we assess the validity of our results using three alternative measures of poverty: (1) the National Center for Education Statistics' (NCES) school neighborhood poverty estimates for the 2020-21 school year (Geverdt & Nixon, 2018), which takes the average income-to-poverty ratio across schools within each district; (2) estimates maintained by SEDA of the share of students qualifying for FRPL for the 2021-22 school year that are corrected for participation in the Community Eligibility Program;¹⁸ and (3) Small Area Income and Poverty Estimates (SAIPE) from the Census Bureau for 2022, which provides the lion's share of Title I formula eligible populations.

We link the CCD data on the universe of public school districts in the U.S. to data that the Department of Education (ED) uses to determine Title I funding. Specifically, the Title I formula used to determine federal allocations to districts is based predominantly on FEC and FEP.¹⁹ The Census' SAIPE data present counts of children between the ages 5 and 17 who are in poverty in a district's geographic area, which are estimated using a combination of Internal Revenue Service (IRS) tax return data from the prior year and 5-year American Community Survey (ACS) data (U.S. Census Bureau, 2023). Because SAIPE data are published years after they are collected, and Title I funding is determined prior to the start of each school year, Title I funding uses the most recent (but still lagged) data available to determine allocations. For the 2020-21 school year, in which ESSER II and ESSER III were determined, ED used 2018 SAIPE

¹⁸ SEDA describe their process for constructing these covariates and correcting for Community Eligibility in detail in their documentation for SEDA 4.1 (Fahle et al., 2021).

¹⁹ Formula-eligible children include the following categories: children between the ages 5 and 17 who are in poverty in a district's geographic area, children qualifying for Temporary Assistance for Needy Families (TANF), neglected and delinquent children, and foster children.

data to determine the number of FEC in the poverty category. We use FEP (predominantly consisting of 2018 children in poverty) to instrument for ESSER spending per pupil.

One weakness of the ESSER *allocations* data is that it is not an appropriate predictor of impacts for a single year of student achievement because any potential impacts would be scaled by the total grant rather than the amount of funds spent in a particular year. For this reason, we attempted to collect district-level spending data for the 2022-23 school year individually from all 30 states in the SEDA 2023 achievement sample. Two states (Alabama and Oklahoma) were unresponsive to or denied these requests despite multiple attempts to obtain the data.

For the 2018-19, 2021-22, and 2022-23 school years, SEDA has published district-by-subgroup achievement data as well as district covariates (Reardon et al., 2024). Due to limitations and sparsity of test score data in the pandemic, we treat the gap between achievement in 2018-19 and 2021-22 as a measure of learning loss experienced during the pandemic and changes from 2021-22 and 2022-23 as a measure of post-pandemic recovery. This follows the convention established by the authors of this dataset (Fahle et al., 2024). These data depend on two main sources: for 2018-19, achievement data is from the EDFacts database; for 2021-22 and 2022-23, the SEDA team collected state-reported data via web scraping and outreach to states.

The SEDA data are limited in several ways that impact the states and districts we can include in our analytic sample. The 2018-19 and 2021-22 dataset excludes 10 states whose data either consisted of too few categories of achievement or is not reported at the necessary level of aggregation. Data from 2022-23 exclude an additional 10 states, mostly due to changes in state tests or proficiency cutoffs. Districts throughout the country may also be left out of the SEDA data due to missing data (i.e., only one of the two considered school years was reported), a change in their Local Education Agency ID (LEAID), if more than 40 percent of their students

took alternative assessments, if the district does not have geographic boundaries (e.g., charters, specialized districts), or if their data are suppressed due to small cell size.

Given the SEDA data restrictions coupled with the ESSER spending data we collected, we include 4,989 districts in math and 4,824 districts in ELA in our analytic samples, representing 51 and 50 percent of nationwide student enrollment, respectively. In Table 1 we present descriptive statistics for districts that are and are not in our analytic sample due to data availability. We weight all averages by student enrollment, so we can interpret these characteristics as those of districts the typical student in our sample attends. Students in our sample, on average, attend districts that received slightly higher ESSER allocations per pupil, spent slightly more of those allocations in 2022-23, and served fewer enrolled students than those missing from the analysis. Students in our sample also attend school with more students identifying as Asian, Multiracial, or White and fewer students identifying as American Native, Hawaiian/Pacific Islander, or Hispanic than students outside of our sample. Students in our sample are also attending districts that are more likely to be in a suburb and are less likely to be in rural areas and towns.

Appendix Table A.2 describes the districts in the sample by quartile of ESSER II & III allocation per pupil. Districts receiving greater ESSER allocations per pupil were, by design, higher-poverty and also experienced larger learning loss between the 2018-19 and 2021-22 school years. For instance, districts in the top quartile of ESSER allocations (receiving an average of over \$4,000 per pupil) had nearly 67 percent of students eligible for FRPL and had learning losses of .18 standard deviations (SDs) in math and .11 SDs in ELA. By contrast, districts in the lowest quartile of ESSER allocations (receiving about \$750 per pupil) had less than 23 percent of students eligible for FRPL and smaller learning losses, about .11 SDs in math

and .07 SDs in ELA. Districts across the distribution of ESSER allocations also differ in terms of urbanicity and student body composition. High allocation districts, for instance, are more likely to be in cities and to serve higher shares of American Native, Black, and Hispanic students.

3.2 *Empirical approach*

As a point of comparison, we begin by estimating the naïve relationship between ESSER spending and achievement using an OLS model:

$$(1) \quad Y_{ij}^{2023} = \beta_0 + \beta_1 ESSER_i + \beta_2 Y_{ij}^{2022} + \beta_3 X_i + \gamma_j + \varepsilon_{ij}$$

Where we predict average achievement in spring 2023 separately for math and ELA (Y_{ij}^{2023}) in district i in state j as a function of the district's ESSER II and III spending per pupil in the 2022-23 school year ($ESSER_i$), a cubic of prior-year achievement in that same subject (Y_{ij}^{2022}), and a suite of district characteristics (X_i), including: district proportions of student racial/ethnic groups, indicators for district urbanicity, and district total revenue per pupil in 2019-20. We also include flags for districts where we replace missing spending values with zero due to the districts having no ESSER funding allocated for ESSER II and ESSER III.

We estimate this model both with and without state fixed effects (γ_j). Our preferred model includes state fixed effects, identifying the effect of ESSER based on within-state variation across districts. The advantage of including state fixed effects is that we do not conflate any state level policies that influence student achievement with ESSER effects. These models will miss ESSER impacts related to state-level variation in ESSER funding, which is driven by formula features such as small state minimum provisions where sparsely populated states receive higher funding per pupil.

Next, we apply the instrumental variable strategy discussed above. Our primary empirical model takes the form of a two-stage least squares regression with the first stage estimating ESSER per pupil spending:

$$(2) \quad ESSER_{ij} = \beta_0 + \beta_1 FEP_i^{2021} + \beta_2 Y_{ij}^{2022} + \beta_3 X_i + \gamma_j + \varepsilon_{ij}$$

We predict ESSER spending in district i in state j as a function of the formula-eligible percentage of children (FEP_i^{2021}) in that district used to determine Title I funding in the 2020-21 school year,²⁰ and all of the covariates we include in model (1) to predict achievement.²¹ Namely, these include a cubic of lagged achievement from 2021-22, the same vector of district characteristics as above (X_i), and state fixed effects for some models (γ_j). We estimate this model separately for the districts for which we observe math and ELA outcomes.

We present the results from this first stage regression in Table 2. The estimated findings suggest that a 10 percentage-point increase in FEP in a district increases the average ESSER II & III spending per pupil by roughly \$6,000. F-tests suggest FEP is a strong instrument, measuring above 210 for both subject samples.

Having assessed the integrity of the instrument, we estimate our two-stage least squares (2SLS) model as follows:

$$(3) \quad Y_{ij}^{2023} = \beta_0 + \beta_1 \widehat{ESSER}_i + \beta_2 Y_{ij}^{2022} + \beta_3 X_i + \gamma_j + \varepsilon_{ij}$$

²⁰ We additionally estimated these models using both FEP and FEC as instruments for ESSER. This did not meaningfully change our results, however the model failed a test for overidentification (Basmann, 1960; Sargan, 1958) and therefore we only present results using FEP alone. It is also worth emphasizing here that while we control for district characteristics such as demographics, these factors are not included in the ESSER formula and thus do not determine allocation amounts.

²¹ We weight the regressions based on district enrollment; the findings are nearly identical if instead we weight by the number of test takers in each district.

Where all variables included are as defined earlier. Note that for all models we use 2023 year-scale achievement in math and reading from SEDA as our outcomes (Y_{ij}^{2023}), controlling for lagged achievement in the same subject in 2021-22.

As we note above, the identifying assumption of this model is that, *once we control for 2022-23 district-level characteristics*, FEP in a district should only affect test score achievement through its determination of ESSER allocations. FRPL eligibility and formula eligibility are governed by slightly different definitions. Specifically, students qualify for FRPL if their household income is lower than 1.85 times the federal poverty line (again, in schools implementing the Community Eligibility Provision all students are counted as FRPL-eligible). Title I formula eligibility, by contrast, is determined by the Census Bureau's estimates of youth in a district's geographic boundary between the ages 5 to 17 whose household income is below the poverty line in addition to children in foster care, those qualifying for Temporary Assistance for Needy Families (TANF), and delinquent children. Figure 3 demonstrates how some states like Mississippi are greatly impacted by Community Eligibility Provision and how the relationship between these measures differs across contexts.

This model leverages three main sources of identifying variation: first, the time lag in reporting means that 2018 FEC and FEP determined funding for students in 2020-21 and may not reflect 2022-23 student characteristics; second, not all children included in estimations of FEC and FEP are enrolled in public schools; third, measurement error in the estimates that comprise FEC and FEP. Of these three, we can only observe the first source of variation. We estimate an alternative specification where we use the change in FEP between 2018 and 2022 as an instrument, controlling for FEP in 2022 instead of FRPL. The F -statistic on the differential FEP instrument is only 6.5, i.e., it is a weak instrument (Lee et al., 2022; Young, 2022). As a

result, our estimates were far less precise, suggesting that it is the combination of these sources rather than the temporal difference alone which identifies our relationship of interest.

4. Results

In Table 3 we present OLS and 2SLS estimates of the effect of increases in school spending on spring 2023 state assessments in math (Panel A) and ELA (Panel B). We show results with and without state fixed effects. We begin by presenting the naïve OLS estimates, which are statistically significant in both math and ELA. Controlling for district demographics, pre-pandemic revenue, and prior achievement, our point estimates suggest that increases of \$1,000 per pupil in ESSER spending in 2022-23 is associated with gains in student math and ELA achievement of .005 and .021 SDs respectively (column 1). When we only consider variation within states (column 2), however, the estimate for math is practically unchanged at .006 SDs, but the ELA estimate attenuates by about 70% to .006.

Columns 3 and 4 present estimates from the 2SLS specifications where we instrument for spending with FEP. For math, the magnitude of the point estimates without state fixed effects are nearly four times as large as the OLS estimates, but there is little change between the OLS and 2SLS models in the specifications that include state fixed effects.²² The results remain statistically significant, though only marginally so when we look within states. In ELA, the 2SLS result without fixed effects is similar to the OLS estimate. The estimated effect of ESSER spending meaningfully attenuates with the inclusion of state fixed effects. In our preferred model (column 4), the point estimate is close to zero and is no longer statistically significant.²³

²² We test whether the OLS and 2SLS coefficients are significantly different from one another and are unable to reject the null hypothesis that the effects are distinguishable (Clogg et al., 1995).

²³ As was the case in math, the ELA results in the 2SLS specifications are not distinguishable from the OLS models.

The contrast between the 2SLS models with and without state fixed effects suggests that state-level differences in ESSER funding, and hence spending, are an important explanation of differences in academic achievement. It may also suggest that other state-level factors correlated with differences in ESSER funding influenced academic recovery. These possibilities are not necessarily mutually exclusive. States were prohibited from mandating how districts spent ESSER funds (U.S. Department of Education, 2022), but they may have still influenced spending decisions, timelines, and recovery through their approval of district spending plans (Lieberman, 2022). States could also have influenced achievement through investments made using the 10% of ESSER funds allocated to them (Council of Chief State School Officers, 2024) or by incentivizing certain types of recovery programs.²⁴

We further explore the implications of state policies and district context (in **Appendix Table A.3**), where we estimate models that include additional controls for the share of district instruction in 2020-21 conducted remotely and in hybrid format and/or various measures of state characteristics and education policy environments. While our estimates are robust to including additional covariates in isolation—for example, adding controls for incidence of COVID-19 infections—when we include the full set of additional state characteristics, our results are meaningfully attenuated and almost identical to our preferred results, which include state fixed effects. This suggests that some of these state-level factors (or correlates with them) explain the larger magnitude of the ESSER coefficients in the non-fixed effects model specifications.²⁵

²⁴ For example, Texas House Bill 4545 “requires Texas school districts to implement a at a minimum supplemental instruction, an accelerated learning committee, and modified teacher assignments” (Texas Education Agency, 2021). Other states have instead used their state-withheld ESSER funds to support their policy priorities, signaling to districts how to use the recovery money. For example, Tennessee spent over \$200 million on summer school and tutoring, Georgia put \$5 million towards engagement and attendance in rural areas, and Connecticut used \$36 million to fund after school and summer programming (Council of Chief State School Officers, 2024).

²⁵ We additionally test a specification where we include the interaction between FEP and state fixed effects in the first stage to allow for variation in the allocations per pupil to vary more flexibly across states, but do not find meaningfully different results from those reported below.

5. Placebo and Robustness Tests

A central concern about our identification strategy is that our findings could be biased if unobserved aspects of school districts are related to our sources of identifying variation (changes in poverty over time and aspects of our measures) and achievement. To address this, we conduct a placebo test and present the results in Table 4.

The placebo test is centered on the idea that differences in ESSER spending should not impact test achievement *in the past*. While ESSER spending should not affect past achievement, Title I spending, which is correlated with ESSER allocations, would be expected to influence achievement. Thus, we estimate models allowing for (non-causal) effects of future ESSER spending on past achievement and compare the magnitudes of the ESSER coefficient estimates to those in our main specifications (in Table 3) and the estimated effects of Title I spending on achievement in a contemporaneous year. Since ESSER spending is roughly four times as large as Title I spending, we might expect the magnitude of estimated coefficients on future ESSER spending to be four times *or less* smaller.²⁶ There are several reasons to think the marginal effects of Title I would be larger than those for ESSER spending: 1) diminished marginal returns to ESSER dollars associated with the larger ESSER allocation; 2) the fact that Title I funding has been a stable, long-term funding source as compared to the time-limited nature of ESSER, which likely presented challenges to implementing academic recovery programs (Carbonari et al., 2022).

²⁶ Note that ESSER *spending* in 2022-23 is not perfectly correlated with Title I spending in 2018-19, the correlation between the two is roughly 0.3. There are two reasons for this. The first is that the pace of spending in 2022-23 varies across school districts and the second is that FEP and FEC counts change over time. Both sources of variation between ESSER and Title I are relevant for the OLS specification of the model, while only the change in FEP and FEC influence the 2SLS specifications (since we are instrumenting for spending in those).

We implement the placebo test by estimating student achievement in a pre-pandemic period, the 2018-19 school year, and using model specifications analogous to those that generate the findings in Table 3, but with modifications. Specifically, in the specifications in columns 1-4, we have lagged all district controls and the instruments for that fiscal year (in columns 3 and 4) but replaced ESSER spending (or the instrumented values for it) with Title I spending (or the instrumental values) in 2018-19.²⁷ This provides baseline estimates of the effects of Title I spending. We find that the impacts of Title I on student achievement are generally larger than those we find for ESSER (in Table 3), particularly in the fixed effects models identified by within state variation in achievement. The estimates of Title I spending are only statistically significant in two models: the OLS estimate for math with state fixed effects and our 2SLS estimate for ELA without fixed effects.

In columns 5 through 8, we replace Title I spending in 2018-19 with the 2022-23 ESSER spending levels.²⁸ In almost all specifications the point estimates are also far smaller than those presented in the analogous specifications reported in Table 3. Given that the relationship between 2023 ESSER spending and district revenues is about four times stronger in 2023 (when federal outlays include both Title I and ESSER) than in 2019 (when it includes only Title I), we should see commensurate increases in the relationship between ESSER spending achievement. Comparing Table 4 to Table 3, the coefficients on ESSER allocations are about a third as large in 2019, which is consistent with the exclusion restriction. That said, this coefficient may also be explained by district poverty predicting other interventions (e.g., in-school attendance)

²⁷ Note that we can only estimate this model for a subsample of districts in our main table due to missingness of prior test scores and FEP.

²⁸ The one change between this and the Title I specification is that we use the same instrument from our main models. This results in our identifying variation depending slightly less on the time lag in FEP reporting relative to funding release than it does in our main model.

implemented during the pandemic. Unfortunately, particularly when we instrument, the estimates are less precise and as a result in the 2SLS model with state fixed effects (our preferred specification), we cannot rule out impacts as large as 0.018 and .009 for math and ELA, respectively. All told, while we lack statistical precision to make strong claims based on these tests, we interpret the results as supportive of the integrity of our identification strategy.

Because our identification strategy also hinges on the inclusion of controls for current district characteristics, and in particular, student poverty, our second test is to assess the robustness of our results to alternative district poverty measures. We present these results in Table 5.²⁹ Our preferred estimates (Table 3, column 4) use the share of students in a district in 2022-23 qualifying for FRPL according to the CCD to capture contemporaneous district poverty. If instead we use the average neighborhood income-to-poverty ratio among schools in each district from NCES (Table 5, column 1),³⁰ our results in both subjects are larger in magnitude, though only slightly so in math. The ELA estimate rises to 0.012 SD gains for a marginal \$1,000 increase in spent ESSER. Second, we use the SEDA-estimated control for the share of students in a district qualifying for free/reduced-price lunch in 2021-22 (sample correlation with CCD FRPL .77), we find estimates in both subjects are larger than our preferred model, again more materially so for ELA, and statistically significant at the 1% level. Finally, if we drop 100% CCD FRPL districts (column 3)—that is, districts where every school qualifies for CEP—our estimates are less precise but almost identical to our preferred model in magnitude.

²⁹ We have also assessed the robustness of our results to controlling for multiple years of lagged scores, controlling for lagged scores as linear terms instead of cubics, and a variety of alternative specifications which are available upon request. Our results are almost identical when we weight by the number of students tested (i.e., how many tests contribute to the average scores) instead of total district enrollment. When we weight our results by the precision of the outcome estimate (i.e., the inverse of the standard error of the achievement outcome) we get slightly larger magnitude of estimated impacts which are more precise. Generally, we find that while our models struggle with precision, our point estimates tend to fall in a somewhat consistent—if not statistically indistinguishable—range from one another.

³⁰ Sample correlation between the NCES income-to-poverty ratio and CCD FRPL is -.59. Note that the NCES income to poverty measure decreases in the case of greater poverty so we expect a negative correlation with FRPL measures.

We interpret these assessments as reinforcing the positive finding of ESSER on math achievement and undermining the robustness of the findings for ELA. Across three alternative measures of poverty, our math estimates still fall within a similar range to our preferred model, suggesting our results are not just a result of the specific differences between the FEP and FRPL measures, alone. For ELA, the .002 SD point estimate of the impact of ESSER on achievement appears sensitive to the poverty control included in the model. While we find a similar estimate when we drop 100% FRPL districts (.001 SD, column 3), our estimates are notably larger and statistically significant with the two preceding alternative poverty measures. On balance, we take these results to suggest that while the impact of ESSER on ELA is likely positive, our results do not suggest a robustness to alternative specifications and our findings on the magnitude of this effect are inconclusive.

6. Heterogeneity

Next, we consider whether our estimates mask differences in the impacts of ESSER funds on achievement across districts and the students they serve. In Figure 5, we report a series of heterogeneity analyses by different measures of poverty. There is a suggestive pattern that students in high-poverty districts benefit more from ESSER than those in lower-poverty settings. In both math and ELA, the ESSER coefficient is largest for districts in the top quartile by FRPL (bottom quartile by NCES income-to-poverty ratio, which is scaled so that higher values represent relative wealth) and decreases for districts with lower levels of poverty. And while the ELA estimates are not statistically significant for the full sample, they are significant for districts in the top two quartiles of poverty. The confidence intervals for the lowest-poverty districts are consistently large because there is only a small amount of variation in ESSER spending to exploit for those districts that received the least money. This result is also echoed when we estimate models on student subgroup-specific average test performance. Looking within states,

we find that students who are economically disadvantaged yielded larger positive gains from ESSER funding than the average student in their district.³¹

While we have additionally assessed our effects for evidence of heterogeneity along the lines of student race representation, district pre-pandemic spending, district pandemic learning loss, district average 2022 test scores, and urbanicity, we only find meaningful patterns by the latter distinction. We present visual representations of these results in Appendix Figures A1 to A4. In Appendix Figure A4, which subdivides districts into groups according to urbanicity, we find positive, statistically significant estimates for all urbanities except for suburban areas.

7. Discussion & Conclusion

In this paper, we investigate the academic impact of ESSER, the largest one-time federal investment in K12 schools. Consistent with the literature on funding impacts (e.g., Jackson & Mackevicius 2024), we find that additional ESSER funding leads to student achievement gains in math, although we do not find consistent results in ELA. We interpret our findings as causal because they are based on plausibly exogenous sources of funding variation and are robust to a variety of specification checks.

Our work contributes to the literature on school funding by using a novel identification strategy that exploits allocation rules associated with federal pandemic relief funding. We isolate variation in ESSER spending driven by historical differences in the share of children in a district's geographic area counted as formula-eligible for Title I and the differences between that estimated measure and observed district poverty to estimate ESSER's impact on achievement.

³¹ SEDA also publishes achievement data for subgroups of students in districts when available. Because subgroups are not consistently reported, we can only estimate models on these outcomes for a subsample of districts. That said, when comparing the impacts of ESSER on test achievement for economically disadvantaged children, the impacts were larger than the estimated impacts on all students in that same subset of districts. These results are available upon request.

This approach allows us to examine the impact of spending across and within states for a larger sample of school districts than is possible in prior studies that use court-mandated policy shifts to identify spending variation.

We estimate that, on average, each \$1,000 increase in ESSER per pupil spending led to statistically significant increases in district math test scores of 0.007 standard deviations and a statistically insignificant, more modest increase in ELA scores. This estimate for math is broadly comparable to the meta-analytic estimates of causal research provided by Jackson and Mackevicius (2024). At the same time, our results may understate ESSER's effects because some ESSER funds were used for physical plant improvements (e.g., HVAC), whose effects could extend beyond the study's timeframe (Biasi et al., 2024). Although districts are far from full recovery, our results suggest ESSER has contributed to addressing learning losses in math.

We put the contribution of ESSER funding into context in Figure 5, which shows the learning loss from 2019 to 2022 and the extent to which there is recovery in the 2022-23 school year. Between 2019 and 2022, student achievement in our sample dropped by about .151 SDs in math and .090 SDs in ELA and recovered by .045 SDs in math and .020 SDs in ELA in the 2022-23 school year. Not all the recovery in the 2022-23 SY was because of ESSER funding, however. We can see this by scaling our preferred estimates (0.007 SDs for math and 0.002 for ELA) by the average ESSER spending per pupil in 2022-23 (\$1,170 for our math sample, \$1,166 for our ELA sample) and then dividing this product by the average growth observed for each sample between spring 2022 and spring 2023 (0.045 for math, 0.020 for ELA). This calculation suggests that about 18 percent of the 2022-23 math recovery and 12 percent of the ELA recovery can be attributed to ESSER III funding.

Our heterogeneity analyses show that the average effects of ESSER also mask some differences across district characteristics, namely district poverty and rurality. Other research shows that low-income students and students of color experienced the greatest losses during the pandemic (Fahle et al., 2023), but we find no consistent evidence of differences in ESSER recovery impacts along these dimensions. We find that ESSER accelerated learning recovery least in suburbs, with no significant impacts, while all other urbanities benefitted in math and ELA. Districts with higher rates of local poverty (typically those that also got the most funding) appeared to also have slightly larger impacts per marginal dollar of ESSER spending, although estimates for lower-poverty districts are imprecise.

There are several limitations to our study to keep in mind. Primarily, as noted earlier, our results are not nationally representative since they are based on data from only 28 states and about half of all students nationwide. Second, our analysis of academic recovery compares different cohorts of students. As a result, we may miss nuances related to any varied impacts of COVID across age groups or the movement of students in and out of testing grades. Third, we focus only on test scores, overlooking non-test outcomes that have long-term implications for postsecondary success (Backes et al., 2022; Jackson, 2018) and that also remain below pre-pandemic averages (Malkus, 2024). When we include ESSER spending in our models, we assume that these dollars did not supplant other district funding. If this were the case, however, we would attribute any observed impact onto more dollars than the actual spending increase, yielding more conservative estimates. Finally, while we can characterize impacts of ESSER broadly, our estimates for ELA appear sensitive to specification.

Pinning down *how* ESSER funding impacted student learning is difficult because data on ESSER spending also has several limitations. Official reports on the use of ESSER funds, for

example, are often vague and school budgets are fungible (Goldhaber et al., 2025). Because the federal government delivered the funds with few strings attached, there are no oversight mechanisms to track their usage and gauge their effectiveness (Roza, 2022). As noted earlier, we know from ESSER spending plans that districts intended to leverage the funds in different ways (Bryant et al., 2022; Dusseault & Pillow, 2021; Malkus, 2021). But because data on *actual* spending are not current and ESSER spending is fungible (Goldhaber et al., 2025), rigorously examining *how* ESSER spending impacted student achievement is challenging.³²

Although ESSER spending had positive effects in math, the trajectory of academic recovery suggests many students have yet to catch up to pre-pandemic levels of achievement, and that ELA performance, in particular, has continued to decline since the end of the pandemic (Dewey et al., 2025; Fahle et al., 2024). Our estimates provide some insight into how much future investment may be needed for full recovery. As shown in Figure 5, the loss from pre-pandemic levels at the end of 2023 school year remained .11 SDs in math and .07 SDs in ELA, since some recovery had already occurred by the end of the 2021-22 school year (Fahle et al., 2024). To recoup the remaining loss in math, our estimates suggest schools would need \$15,380 additional funds per pupil, assuming the return on those funds is similar to what we estimated for prior ESSER spending.³³ Scaling this by roughly 50 million public school students yields about \$770 billion. These estimates do not preclude the existence of other resources—including the potential that student recovery continues to exceed what is explained by ESSER impacts alone. Still, the estimated cost of recovery is striking. To put it in perspective, public schools spent an

³² State officials we spoke to while collecting ESSER aggregate spending data for 2022-23 voiced concerns about the accuracy and quality of any *disaggregation* of these totals across expense types, as the timelines and categorizations required for reporting were in flux throughout the grant period. Additionally, the fungibility of these dollars across line items and the remainder of the district budget makes the quality and reliability of fully itemized data dubious.

³³ This is calculated by taking the .15 loss in math and dividing by the ESSER coefficient in Table 3, column 4 of Panel A (0.007).

average of about \$14,800 per pupil in the 2019-20 school year (National Center for Education Statistics, 2024), so our estimate is more than double current per pupil spending levels. Our analysis suggests that ESSER funding helped address some of the academic decline students experienced during the pandemic. But, as others have noted, the amount of funding schools have is not all that matters; *how* schools use funding—in both purpose and efficiency—also matters (Handel & Hanushek, 2023; McGee, 2023).

While this extrapolated cost of recovery may appear high, estimates of leaving pandemic-related losses unaddressed suggest that large additional expenditures may be worthwhile. Doty et al. (2022) argue that the decline in test scores from the pandemic could translate into \$900 billion in future foregone wages. Hanushek and Strauss (2024) predict an even higher cost of unremedied learning loss: \$31 trillion, about 35 times Doty et al.’s (2022) estimate.³⁴ In other words, while the immediate costs seem remarkable, they represent an investment in youth that will have positive impacts multiple times over as they develop and, eventually, enter the workforce.

³⁴ The difference between the two estimates stems from the fact that Doty and others consider lost wages alone, while Hanushek and Strauss consider lost wages as well as broader effects on the nation’s productivity and growth. Regardless of which assumptions are privileged, both estimates far exceed the full-recovery cost estimate.

References

- American Enterprise Institute. (2021). *Return to Learn Tracker Data* (Version 45e) [Data set]. https://www.returntolearntracker.net/instructional_status/
- Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2022). *Teachers and School Climate: Effects on Student Outcomes and Academic Disparities*. https://caldercenter.org/sites/default/files/CALDER%20Working%20Paper%20274-1022_0.pdf
- Baron, E. J., Hyman, J., & Vasquez, B. (2024). Public School Funding, School Quality, and Adult Crime. *The Review of Economics and Statistics*, 1–46. https://doi.org/10.1162/rest_a_01452
- Basmann, R. L. (1960). On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics. *Journal of the American Statistical Association*, 55(292), 650–659. <https://doi.org/10.2307/2281588>
- Biasi, B., Lafortune, J. M., & Schönholzer, D. (2024). *What Works and For Whom? Effectiveness and Efficiency of School Capital Investments Across The U.S.* (Working Paper No. 32040). National Bureau of Economic Research. <https://doi.org/10.3386/w32040>
- Borman, G. D., & D’Agostino, J. V. (1996). Title I and Student Achievement: A Meta-Analysis of Federal Evaluation Results. *Educational Evaluation and Policy Analysis*, 18(4), 309–326. <https://doi.org/10.2307/1164335>
- Bryant, J., Dorn, E., Redmond, W., Shelat, N., & Williams, C. (2022). *Halftime for the K-12 stimulus: How are districts faring?* McKinsey & Company. <https://www.mckinsey.com/industries/education/our-insights/halftime-for-the-k-12-stimulus-how-are-districts-faring#/>
- Maria V. Carbonari, Miles Davison, Michael DeArmond, Daniel Dewey, Elise Dizon-Ross, Dan Goldhaber, Ayesha Hashim, Thomas J. Kane, Andrew McEachin, Emily Morton, Tyler Patterson, Douglas O. Staiger (2022). The Challenges of Implementing Academic COVID Recovery Interventions: Evidence from the Road to Recovery Project. CALDER Working Paper No. 275-1222.
- Cascio, E. U., Gordon, N., & Reber, S. (2013). Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South. *American Economic Journal: Economic Policy*, 5(3), 126–159. <https://doi.org/10.1257/pol.5.3.126>
- Chen, J. T., & Krieger, N. (2021). Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County Versus Zip Code Analyses. *Journal of Public Health Management and Practice*, 27(Supplement 1), S43. <https://doi.org/10.1097/PHH.0000000000001263>

- Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5), 1261–1293.
- Council of Chief State School Officers. (2022). *The road to recovery: How states are using ESSER funds to support student and staff mental health and wellbeing*. <https://learning.ccsso.org/states-leading-how-state-education-agencies-are-leveraging-the-esser-set-aside>
- Council of Chief State School Officers. (2024). *States Leading: How State Education Agencies are Leveraging the ESSER Set-Aside*. <https://learning.ccsso.org/states-leading-how-state-education-agencies-are-leveraging-the-esser-set-aside>
- Dee, T. S. (2024). Higher chronic absenteeism threatens academic recovery from the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 121(3), e2312249121. <https://doi.org/10.1073/pnas.2312249121>
- Dewey, D. C., Fahle, E. M., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2024). *Federal Pandemic Relief and Academic Recovery* (Working Paper No. 32897). National Bureau of Economic Research. <https://doi.org/10.3386/w32897>
- Dewey, D. C., Fahle, E. M., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2025). *Pivoting from Pandemic recovery to long-term reform: A district-level analysis*. Education Recovery Scorecard. <https://educationrecoveryscorecard.org/wp-content/uploads/2025/02/Pivoting-from-Pandemic-Recovery-to-Long-Term-Reform-A-District-Level-Analysis.pdf>
- Doty, E., Kane, T. J., Patterson, T., & Staiger, D. O. (2022). *What Do Changes in State Test Scores Imply for Later Life Outcomes?* (Working Paper No. 30701). National Bureau of Economic Research. <https://doi.org/10.3386/w30701>
- Dusseault, B., & Pillow, T. (2021). *First look at ESSER priorities: Districts are placing their bets on what they know*. Center on Reinventing Public Education. <https://crpe.org/first-look-at-esser-priorities-districts-are-placing-their-bets-on-what-they-know/>
- Fahle, E. M., Chavez, B., Kalogrides, D., Shear, B. R., Reardon, S. F., & Ho, A. D. (2021). *Stanford Education Data Archive, Technical Documentation Version 4.1*. Stanford Education Data Archive. https://stacks.stanford.edu/file/druid:xv742vh9296/seda_documentation_4.1.pdf
- Fahle, E. M., Kane, T. J., Patterson, T., Staiger, D. O., & Stuart, E. A. (2023). *School district and community factors associated with learning loss during the COVID-19 pandemic*. Center for Education Policy Research at Harvard University.
- Fahle, E. M., Kane, T. L., Reardon, S. F., & Staiger, D. O. (2024). *The First Year of Pandemic Recovery: A District-Level Analysis* (Education Recovery Scorecard). Center for Education Policy Research, Harvard University.

- Finch, W. H., & Hernández Finch, M. E. (2020). Poverty and Covid-19: Rates of Incidence and Deaths in the United States During the First 10 Weeks of the Pandemic. *Frontiers in Sociology*, 5. <https://doi.org/10.3389/fsoc.2020.00047>
- Geverdt, D., & Nixon, L. (2018). *Sidestepping the box: Designing a supplemental poverty indicator for school neighborhoods* (No. 2017–039). U.S. Department of Education, National Center for Education Statistics.
- Goldhaber, D., Falken, G. T., & Theobald, R. (2025). *ESSER funding and school system jobs: Evidence from job posting data* (No. 297-0225–2; CALDER Working Paper). Center for Analysis of Longitudinal Data in Education Research at the American Institutes for Research. <https://caldercenter.org/publications/esser-funding-and-school-system-jobs-evidence-job-posting-data>
- Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2023). The educational consequences of remote and hybrid instruction during the pandemic. *American Economic Review: Insights*, 5(3), 377–392. <https://doi.org/10.1257/aeri.20220180>
- Gordon, N. (2004). Do federal grants boost school spending? Evidence from Title I. *Journal of Public Economics*, 88(9), 1771–1792. <https://doi.org/10.1016/j.jpubeco.2003.09.002>
- Gordon, N., & Reber, S. (2023). *Title I of ESEA: How the formulas work* (Understanding and Improving Title I of ESEA). All4Ed. <https://all4ed.org/publication/title-i-of-esea-how-the-formulas-work/>
- Hall, L. R., Sanchez, K., da Graca, B., Bennett, M. M., Powers, M., & Warren, A. M. (2022). Income Differences and COVID-19: Impact on Daily Life and Mental Health. *Population Health Management*, 25(3), 384–391. <https://doi.org/10.1089/pop.2021.0214>
- Handel, D. V., & Hanushek, E. A. (2023). Chapter 3 - US school finance: Resources and outcomes. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 7, pp. 143–226). Elsevier. <https://doi.org/10.1016/bs.hesedu.2023.03.003>
- Hanushek, E. A. (1989). The Impact of Differential Expenditures on School Performance. *Educational Researcher*, 18(4), 45–62. <https://doi.org/10.3102/0013189X018004045>
- Hanushek, E. A. . (1994). Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald. *Educational Researcher*, 23(4), 5–8. <https://doi.org/10.3102/0013189X023004005>
- Hanushek, E. A. ., & Strauss, B. (2024). *A global perspective on US learning losses*. Hoover Institution. https://www.hoover.org/sites/default/files/research/docs/Hanushek-Strauss_WebreadyPDF_240229.pdf

- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). An Exchange: Part I: Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Educational Researcher*, 23(3), 5–14. <https://doi.org/10.2307/1177220>
- Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, 126(5), 2072–2107. <https://doi.org/10.1086/699018>
- Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, 131(1), 157–218. <https://doi.org/10.1093/qje/qjv036>
- Jackson, C. K., & Mackevicius, C. L. (2024). What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States. *American Economic Journal: Applied Economics*, 16(1), 412–446. <https://doi.org/10.1257/app.20220279>
- Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do School Spending Cuts Matter? Evidence from the Great Recession. *American Economic Journal: Economic Policy*, 13(2), 304–335. <https://doi.org/10.1257/pol.20180674>
- Johnson, R. C. (2015). Follow the Money: School Spending from Title I to Adult Earnings. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 1(3), 50–76.
- Karmakar, M., Lantz, P. M., & Tipirneni, R. (2021). Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. *JAMA Network Open*, 4(1), e2036462. <https://doi.org/10.1001/jamanetworkopen.2020.36462>
- Kreisman, D., & Steinberg, M. P. (2019). The effect of increased funding on student achievement: Evidence from Texas’s small district adjustment. *Journal of Public Economics*, 176, 118–141. <https://doi.org/10.1016/j.jpubeco.2019.04.003>
- Kuhfeld, M., Soland, J., Lewis, K., Ruzek, E., & Johnson, A. (2022). The COVID-19 School Year: Learning and Recovery Across 2020-2021. *AERA Open*, 8, 23328584221099306. <https://doi.org/10.1177/23328584221099306>
- Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics*, 10(2), 1–26. <https://doi.org/10.1257/app.20160567>
- Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-Ratio Inference for IV. *American Economic Review*, 112(10), 3260–3290. <https://doi.org/10.1257/aer.20211063>
- Lieberman, M. (2022, May 26). States Grill Districts on How They’ll Spend Their ESSER Money for Learning Recovery. *Education Week*. <https://www.edweek.org/leadership/states-grill-districts-on-how-theyll-spend-their-esser-money-for-learning-recovery/2022/05>

- Malkus, N. (2021). *The \$200 billion question: How much of federal COVID-19 relief funding for schools will go to COVID-19 relief?* American Enterprise Institute. <https://aei.org/wp-content/uploads/2021/08/The-200-Billion-Question.pdf>
- Malkus, N. (2024, January 31). Long COVID for Public Schools: Chronic Absenteeism Before and After the Pandemic. *American Enterprise Institute - AEI*. <https://www.aei.org/research-products/report/long-covid-for-public-schools-chronic-absenteeism-before-and-after-the-pandemic/>
- Matsudaira, J. D., Hosek, A., & Walsh, E. (2012). An integrated assessment of the effects of Title I on school behavior, resources, and student achievement. *Economics of Education Review*, 31(3), 1–14. <https://doi.org/10.1016/j.econedurev.2012.01.002>
- McGee, J. B. (2023). Yes, money matters, but the details can make all the difference. *Journal of Policy Analysis and Management*, 42(4), 1125–1132. <https://doi.org/10.1002/pam.22519>
- National Center for Education Statistics. (2020, August 5). *Understanding School Lunch Eligibility in the Common Core of Data*. NCES Blog. <https://nces.ed.gov/blogs/nces/post/understanding-school-lunch-eligibility-in-the-common-core-of-data>
- National Center for Education Statistics. (2022a). *2022 NAEP mathematics assessments: Highlighted results at grades 4 and 8 for the nation, states, and districts* (NCES 2022124). U.S. Department of Education. <https://www.nationsreportcard.gov/highlights/mathematics/2022/>
- National Center for Education Statistics. (2022b). *2022 NAEP reading assessment: Highlighted results at grades 4 and 8 for the nation, states, and districts* (NCES 2022126). U.S. Department of Education. <https://www.nationsreportcard.gov/highlights/reading/2022/>
- National Center for Education Statistics. (2024). *Public School Expenditures* (Condition of Education). U.S. Department of Education, Institute of Education Sciences. <https://nces.ed.gov/programs/coe/indicator/cmb>
- National Center for Education Statistics. (2025, January 29). Fourth Graders’ Mathematics Scores Improve Following Historic Drops, Eighth Graders Show No Change. *National Assessment of Educational Progress: Press Release*. https://www.nationsreportcard.gov/reports/mathematics/2024/g4_8/supporting-files/2024rm-press-release.docx
- Office of Elementary & Secondary Education. (2021). *U.S. Department of Education Fact Sheet: American Rescue Plan Act of 2021 Elementary and Secondary School Emergency Relief Fund (ARP ESSER)*. https://oese.ed.gov/files/2021/03/FINAL_ARP-ESSER-FACT-SHEET.pdf
- Office of Elementary & Secondary Education. (2024, May 7). *Elementary and Secondary School Emergency Relief Fund*. Office of Elementary & Secondary Education.

- <https://oese.ed.gov/offices/education-stabilization-fund/elementary-secondary-school-emergency-relief-fund/>
- Reardon, S. F., Fahle, E. M., Ho, A. D., Min, J., Kalogrides, D., & Kane, T. J. (2024). *Stanford Education Data Archive* (SEDA 2023) [Data set]. <https://purl.stanford.edu/xt779fj2637>
- Roza, M. (2022, February 3). *Let Student Progress Be the North Star for School Relief Fund Spending*. <https://www.the74million.org/article/roza-without-clear-rules-theres-no-way-to-judge-how-school-relief-funds-are-being-spent-setting-student-progress-as-a-north-star-would-be-a-game-changer/>
- Sargan, J. D. (1958). The Estimation of Economic Relationships using Instrumental Variables. *Econometrica*, 26(3), 393–415. <https://doi.org/10.2307/1907619>
- Stanford, L. (2023, August 1). Voters Don't Think ESSER Funds Have Improved Their Schools, Poll Finds. *Education Week*. <https://www.edweek.org/policy-politics/voters-dont-think-esser-funds-have-improved-their-schools-poll-finds/2023/08>
- Tang, S., Horter, L., Bosh, K., Kassem, A. M., Kahn, E. B., Ricaldi, J. N., Pao, L. Z., Kang, G. J., Singleton, C.-M., Liu, T., Thomas, I., & Rao, C. Y. (2022). Change in unemployment by social vulnerability among United States counties with rapid increases in COVID-19 incidence—July 1–October 31, 2020. *PLOS ONE*, 17(4), e0265888. <https://doi.org/10.1371/journal.pone.0265888>
- Texas Education Agency. (2021). *District Fact Sheet: House Bill 4545 Overview*. <https://tea.texas.gov/texas-schools/health-safety-discipline/covid/hb-4545-overview.pdf>
- U.S. Census Bureau. (2023, November). *2010 - 2022 Overview of School District Estimates*. Census.Gov. <https://www.census.gov/programs-surveys/saipe/technical-documentation/methodology/school-districts/overview-school-district.html>
- U.S. Department of Education. (2022). *Frequently asked questions: Elementary and Secondary School Emergency Relief Programs Governor's Emergency Education Relief Programs*. U.S. Department of Education.
- van der Klaauw, W. (2008). Breaking the link between poverty and low student achievement: An evaluation of Title I. *Journal of Econometrics*, 142(2), 731–756. <https://doi.org/10.1016/j.jeconom.2007.05.007>
- Weinstein, M. G., Stiefel, L., Schwartz, A. E., & Chalico, L. (2009). Does Title I Increase Spending and Improve Performance? Evidence from New York City. Working Paper #09-09. In *Institute for Education and Social Policy*. Institute for Education and Social Policy. <https://eric.ed.gov/?id=ED556781>
- Wolf, E. R., Nguyen, M., Sabo, R. T., Foster, R., Avula, D., Gilbert, J., Freymiller, C., Nelson, B. B., & Krist, A. H. (2024). School Closure and Child Maltreatment During the COVID-19 Pandemic. *Child Maltreatment*, 10775595241252350. <https://doi.org/10.1177/10775595241252350>

Young, A. (2022). Consistency without Inference: Instrumental Variables in Practical Application. *European Economic Review*, 147, 104112.
<https://doi.org/10.1016/j.euroecorev.2022.104112>

Tables and Figures

Table 1. Characteristics of districts for the average student by sample inclusion

	All districts	Math sample	Not in math sample	ELA sample	Not in ELA sample
ESSER II & III allocation per pupil (\$1,000s)	3.0	3.1***	2.8	3.1***	2.9
ESSER II & III spent per pupil 2022-23 (\$1,000s)	1.2	1.2***	1.0	1.2*	1.1
District enrollment (1,000s)	39.2	37.8**	40.8	38.1*	40.4
% Free/reduced price lunch 2022-23	49.7	48.5***	51.0	48.4***	51.0
% American Native 2022-23	0.9	0.4***	1.4	0.4***	1.4
% Asian 2022-23	5.5	6.2***	4.8	6.3***	4.7
% Black 2022-23	14.2	14.3	14.2	14.4	14.1
% Hawaiian/Pacific Islander 2022-23	0.4	0.3***	0.4	0.3***	0.4
% Hispanic 2022-23	27.8	24.9***	30.9	25.0***	30.6
% Multiracial 2022-23	4.9	5.2***	4.6	5.2***	4.6
% White 2022-23	45.3	46.9***	43.7	46.5***	44.1
% Districts in rural area	16.2	15.4**	16.9	15.0***	17.3
% Districts in cities	30.3	30.4	30.1	30.0	30.5
% Districts in suburbs	42.4	43.5***	41.2	44.6***	40.2
% Districts in towns	11.2	10.6**	11.7	10.4***	12.0
District total revenue per pupil 2019-20 (\$1,000s)	15.5	15.5	15.4	15.6**	15.3
<i>N</i>	13,234	4,989	8,245	4,824	8,410

Note. Means are weighted by district enrollment. Stars indicate statistically significant differences from the subject's sample of districts, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. ELA=English Language Arts; ESSER=Elementary and Secondary School Relief Fund.

Table 2. First stage regressions, predicting 2022-23 ESSER II and III spending per pupil (\$1,000s)

	Math sample		ELA sample	
	(1)	(2)	(3)	(4)
Formula-eligible percent (FEP)	5.80*** (0.48)	6.19*** (0.40)	5.67*** (0.57)	6.01*** (0.41)
State fixed effects		X		X
<i>n</i>	4,989	4,989	4,824	4,824
<i>R</i> ²	0.50	0.70	0.50	0.70
Instrument <i>F</i> -statistic	143.36	236.78	97.29	210.29

Note. Heteroskedasticity-robust standard errors are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All models are weighted by district enrollment in 2022-23. *F*-statistics for instruments in each model are listed at the bottom of the table. We report first stages separately for districts for which we observe math and English language arts scores because some states only report data usable by the Stanford Education Data Archive (SEDA) for a single subject. All models include covariates of district characteristics from the 2022-23 school year, including the following: urbanicity indicators, enrolled students race representation and rate of free/reduced-price lunch qualification, lagged average test scores for the sample subject (cubic), and total revenue per pupil in 2019-20. ELA=English language arts; ESSER=Elementary and Secondary School Relief Fund.

Table 3. Estimated impacts of ESSER II and III spending on district-level student achievement

	(1)	(2)	(3)	(4)
<i>Panel A. Math achievement</i>				
ESSER per pupil spending	0.005***	0.006**	0.019***	0.007*
2022-23 (\$1,000s)	(0.002)	(0.003)	(0.005)	(0.004)
<i>n</i>	4,989	4,989	4,989	4,989
<i>Panel B. English language arts achievement</i>				
ESSER per pupil spending	0.021***	0.006***	0.019***	0.002
2022-23 (\$1,000s)	(0.004)	(0.002)	(0.004)	(0.003)
<i>n</i>	4,824	4,824	4,824	4,824
Estimator	OLS	OLS	2SLS	2SLS
State fixed effects		X		X

Note. Heteroskedasticity-robust standard errors are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Models (2) and (4) use district formula-eligible percentage for fiscal year 2020 to instrument for district ESSER II and III 2022-23 per pupil spending. All models are weighted by district enrollment in 2022-23. Models include covariates of district characteristics from the 2022-23 school year, including the following: urbanicity indicators, enrolled students race representation and rate of free/reduced-price lunch qualification, a cubic of 2022 district test scores for the same subject as the outcome, and total revenue per pupil in 2019-20. ESSER=Elementary and Secondary School Relief Fund; OLS=ordinary least squares; 2SLS=two-stage least squares.

Table 4. Placebo test assessing ESSER impact on 2018-19 student achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Math achievement 2018-19</i>								
Title I per pupil 2018-19 (\$1,000s)	0.015 (0.010)	0.023** (0.008)	0.006 (0.015)	0.024 (0.016)				
ESSER per pupil spending 2022-23 (\$1,000s)					0.003 (0.002)	0.002 (0.001)	0.003 (0.005)	0.008 (0.005)
<i>n</i>	4,866	4,866	4,866	4,866	4,866	4,866	4,866	4,866
<i>Panel B. English language arts achievement 2018-19</i>								
Title I per pupil 2018-19 (\$1,000s)	0.022* (0.010)	0.015 (0.008)	0.019 (0.014)	0.005 (0.013)				
ESSER per pupil spending 2022-23 (\$1,000s)					0.005** (0.002)	0.001 (0.001)	0.004 (0.006)	-0.001 (0.005)
<i>n</i>	4,794	4,794	4,794	4,794	4,794	4,794	4,794	4,794
Estimator	OLS	OLS	2SLS FEP (FY18)	2SLS FEP (FY18)	OLS	OLS	2SLS FEP (FY20)	2SLS FEP (FY20)
Instrument								
State fixed effects		X		X		X		X

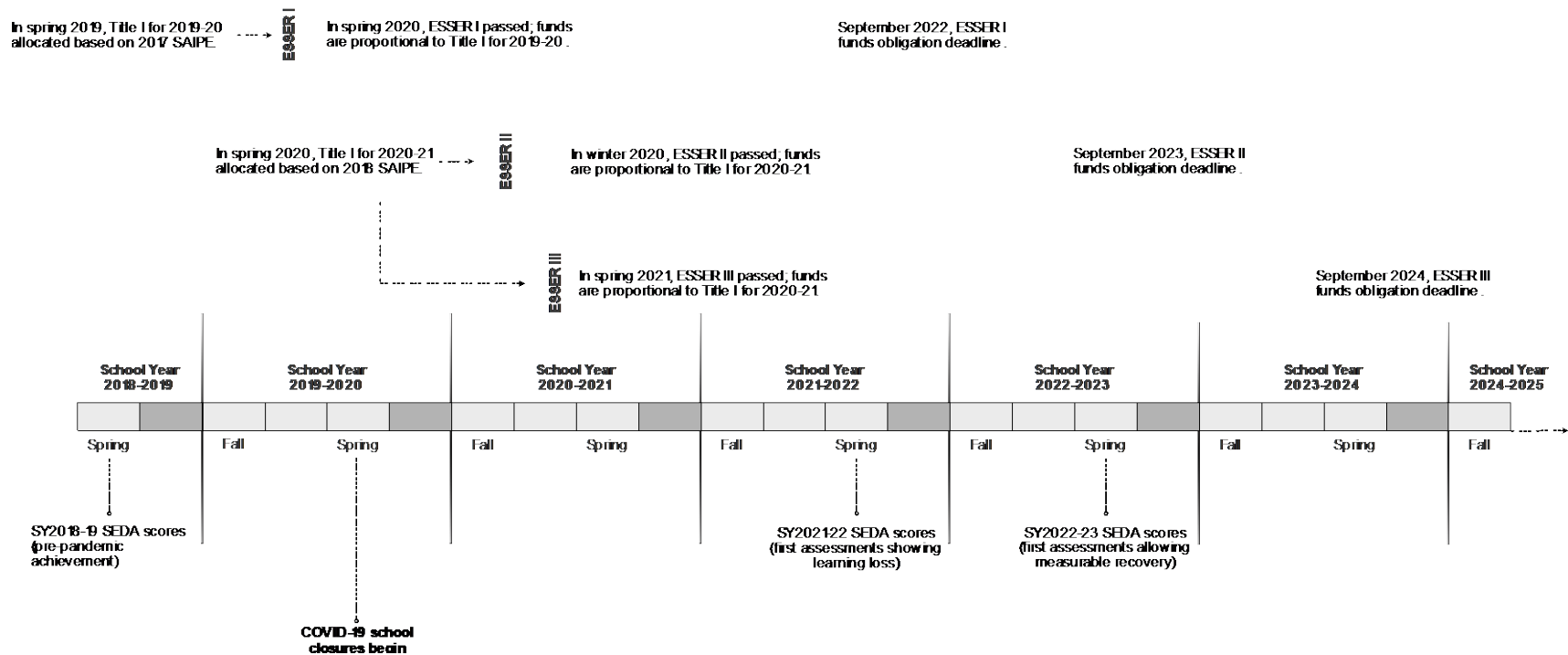
Notes. Heteroskedasticity-robust standard errors are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Models (3) and (4) use district formula-eligible percentage for fiscal year 2018 to instrument for district Title I revenue per pupil in 2018-19. Models (7) and (8) use district formula-eligible percentage for fiscal year 2020 to instrument for district ESSER II and III 2022-23 per pupil spending. All models are weighted by district enrollment in 2022-23. Models include covariates of district characteristics from the 2018-19 school year, including the following: urbanicity indicators, enrolled students race representation and rate of free/reduced-price lunch qualification, a cubic of 2017-18 district test scores for the same subject as the outcome, and total revenue per pupil in 2019-20. ESSER=Elementary and Secondary School Education Recovery Fund; FEP = formula-eligible percentage; FY=Fiscal year; OLS=ordinary least squares; 2SLS=two-stage least squares.

Table 5. Two-stage least squares estimates using alternative poverty measures as controls

	(1)	(2)	(3)
<i>Panel A. Math achievement</i>			
ESSER per pupil spending 2022-23 (\$1,000s)	0.010*	0.011***	0.007
	(0.006)	(0.003)	(0.004)
<i>n</i>	4,989	4,989	4,864
<i>Panel B. English language arts achievement</i>			
ESSER per pupil spending 2022-23 (\$1,000s)	0.012***	0.009***	0.001
	(0.003)	(0.003)	(0.004)
<i>n</i>	4,824	4,824	4,724
Estimator	2SLS	2SLS	2SLS
Poverty measure	NCES IPR	SEDA FRPL	<100% FRPL
State fixed effects	X	X	X

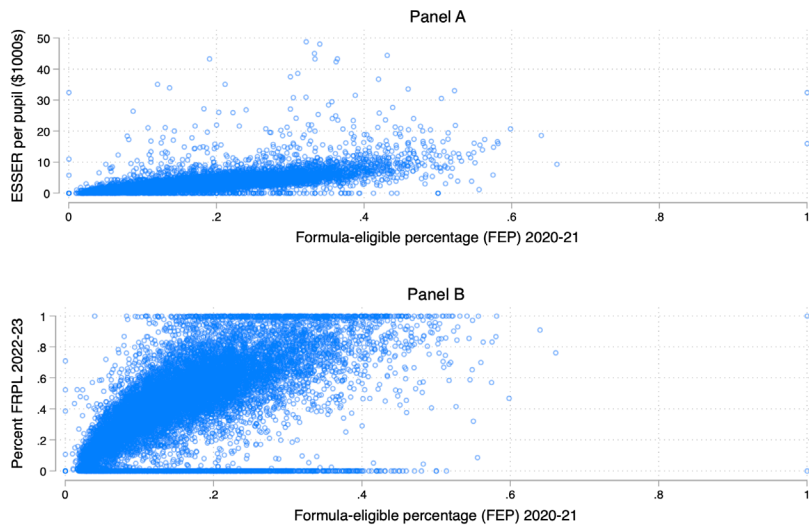
Note. Heteroskedasticity-robust standard errors are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All models use district formula-eligible percentage for fiscal year 2020 to instrument for district ESSER II and III 2022-23 per pupil spending and are weighted by district enrollment in 2022-23. Models include covariates of district characteristics from the 2022-23 school year, including the following: urbanicity indicators, enrolled students race representation, a cubic of 2022 district test scores for the same subject as the outcome, and total revenue per pupil in 2019-20. Each model also includes a unique control for district poverty. Model (1) controls for the average NCES IPR in the district. Model (2) uses SEDA-estimated district FRPL. Model (3) uses FRPL from the Common Core of Data but drops districts with 100% reported FRPL. ESSER=Elementary and Secondary School Relief Fund; FRPL=free/reduced-price lunch; IPR=Income-to-Poverty Ratio; NCES=National Center for Education Statistics; SAIPE=Small-Area Income and Poverty Estimates; SEDA=Stanford Education Data Archive.

Figure 1. Timeline of outcome measures and ESSER events



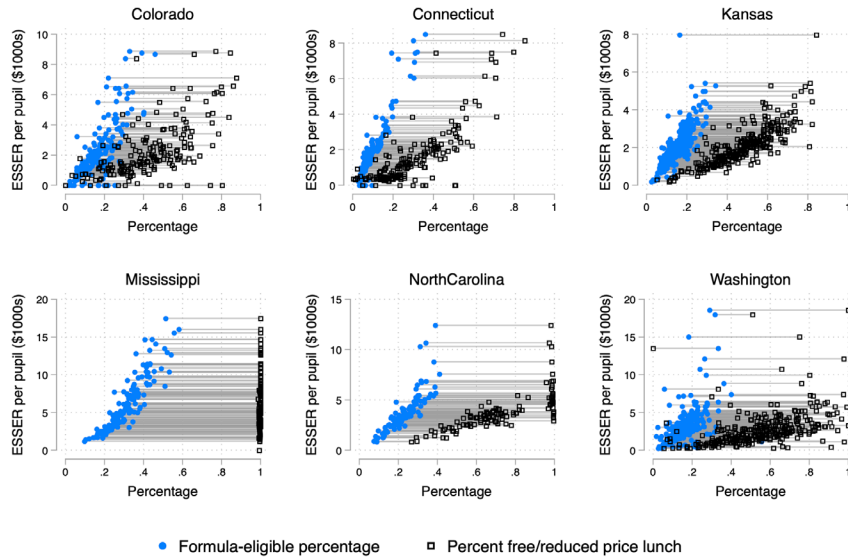
Notes: ESSER I was passed into law in March of 2020; ESSER II was passed in December 2020; ESSER III was passed in March of 2021. Fund dispersals would have been slightly delayed from these dates of passage as they were administered. Title I is largely determined by the number of formula-eligible children in a district, the primary source for which is the SAIPE. ESSER=Elementary and Secondary School Emergency Relief Fund; SAIPE=Small Area Income and Poverty Estimates; SEDA=Stanford Education Data Archive.

Figure 2. Differences in the relationships between ESSER allocations, formula-eligible percents, and free/reduced-price lunch prevalence



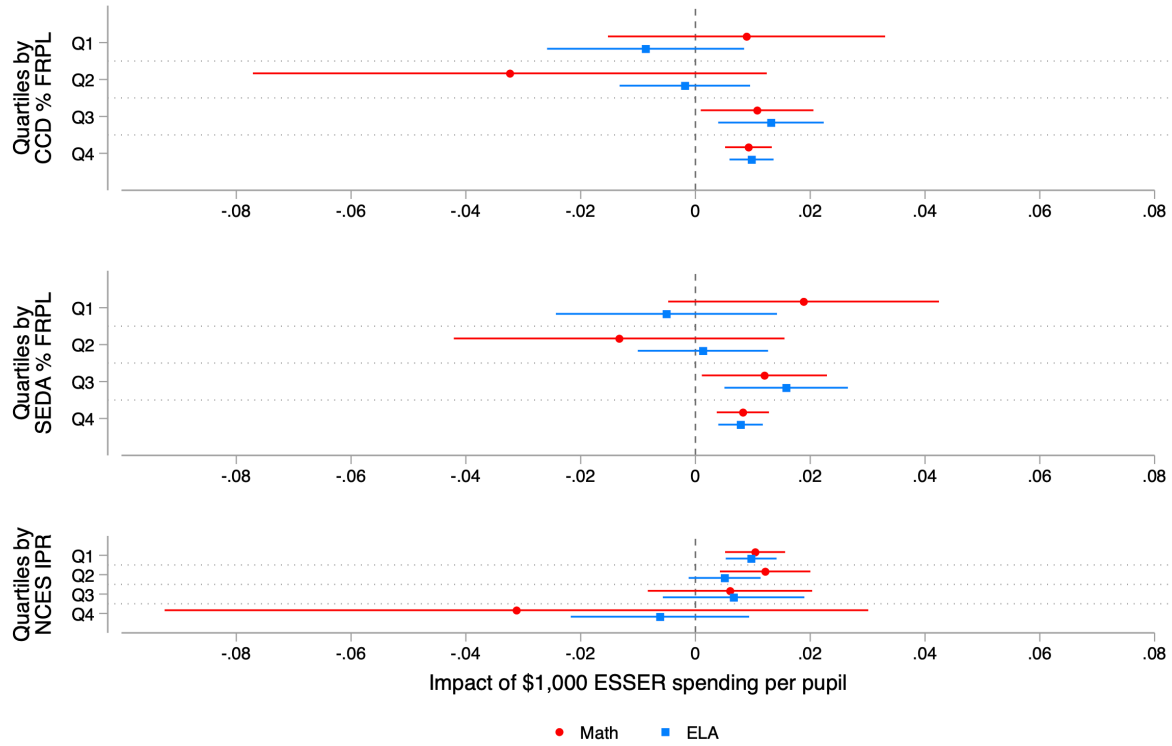
Note. Panel A presents a district-level plot of the relationship between formula-eligible percentages (FEP) for fiscal year 2020 (2020-21 school year) and ESSER II & III allocations per pupil in \$1,000s. Panel B presents FEP and percent free or reduced-price lunch in 2022-23.

Figure 3. Differences in formula-eligible percentage and free/reduced-price lunch across states



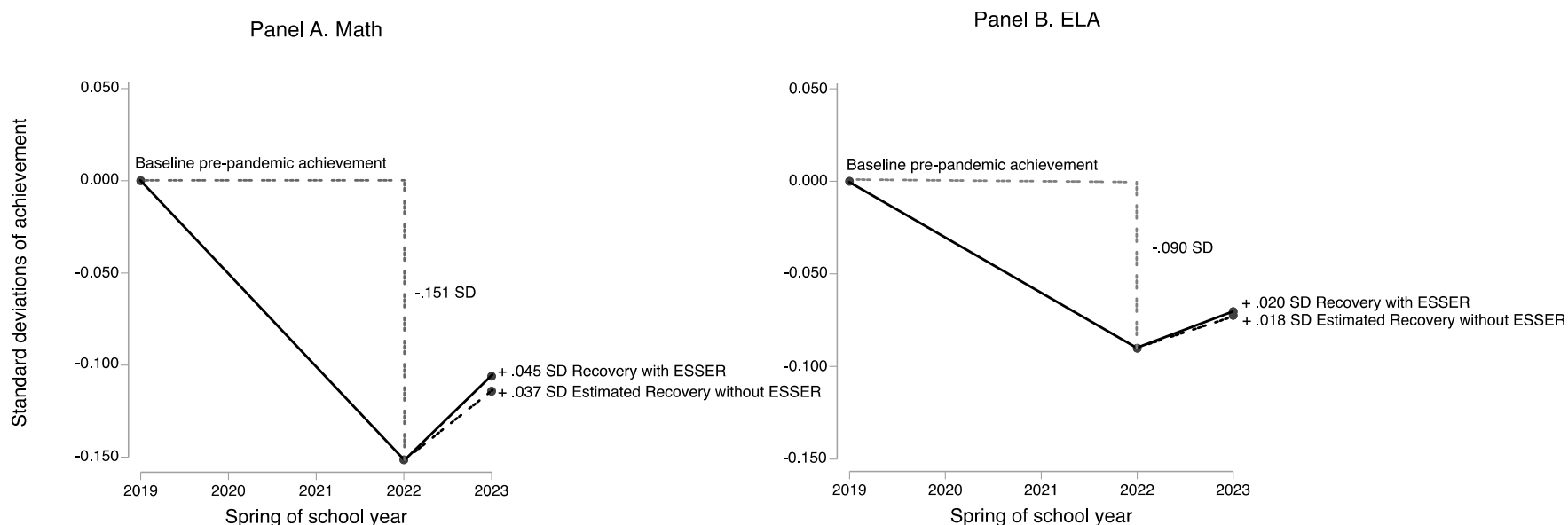
Note. Each dot represents the district value for the indicated measure, with districts connected by the grey lines. Free/reduced price lunch values are for the 2022-23 school year. Formula-eligible percentage uses Small Area Income and Poverty Estimates (SAIPE) data from 2018 and determined Title I allocations for the 2020-21 school year (determined in the 2020 fiscal year). Y-axis ESSER per pupil values are summed across ESSER II and III.

Figure 4. Estimate heterogeneity across quartiles of district poverty



Note. Points represent estimates from our 2SLS model with state fixed effects, estimating on the subsample of districts grouped into quartiles according to the y-axis poverty measure indicated. Spikes present 95% confidence intervals.

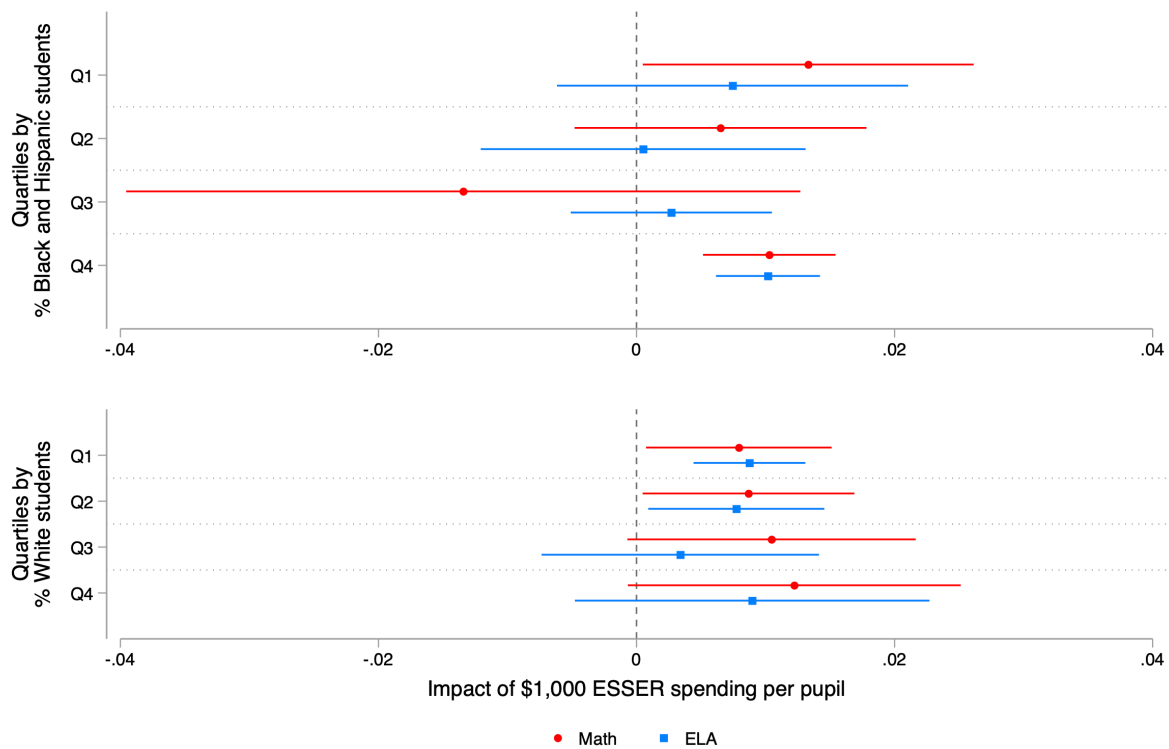
Figure 5. Observed changes in achievement over time and estimated impact of ESSER III funding



Note. Points represent means of district-level achievement weighted by student enrollment; we normalize 2019 achievement to zero for ease of comparison. Gaps in achievement between spring 2019 and spring 2022 capture pandemic learning loss; growth between 2022 and 2023 illustrate post-pandemic recovery. Achievement is standardized by subject and grade such that one unit is one standard deviation. The dashed line in each panel represents our estimated level of achievement without the effect of ESSER III funds we observe. We calculate these values by subtracting the product of our estimates in column 4 of Table 3 by the average ESSER III allocation for each subject sample from the weighted average of 2023 student achievement.

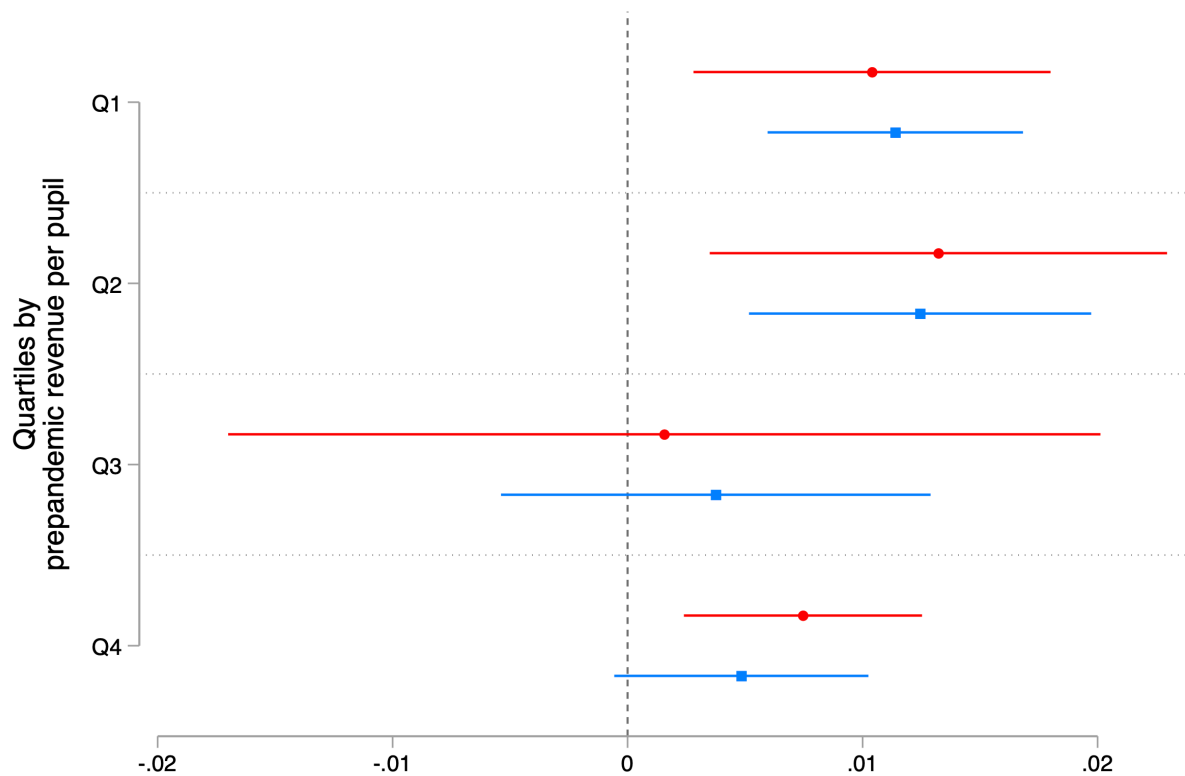
Appendix: Supplemental Results, Figures, and Tables

Figure A1. Estimates of heterogeneity across quartiles of student race/ethnicity



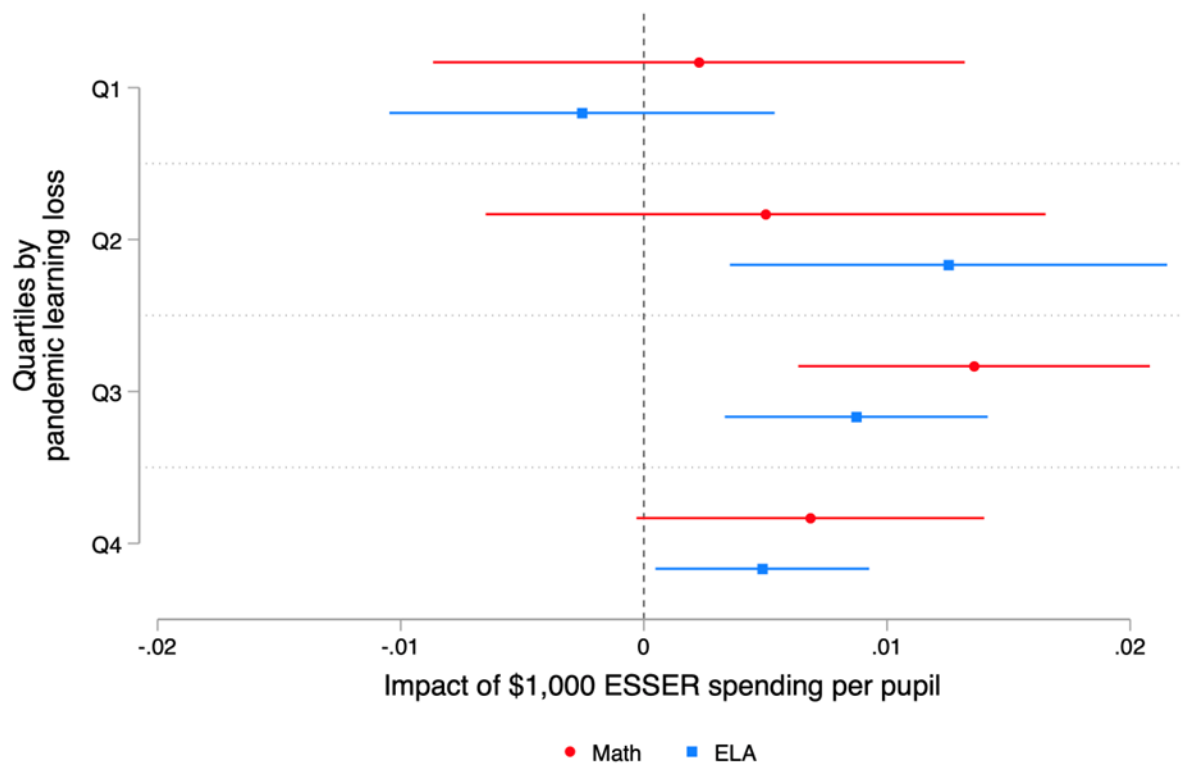
Note. Points represent estimates from our 2SLS model with state fixed effects, estimating on the subsample of districts grouped into quartiles according to the percent of enrolled students in the racial group indicated. Spikes present 95% confidence intervals.

Figure A2. Estimates of heterogeneity across quartiles of prepandemic revenue per pupil



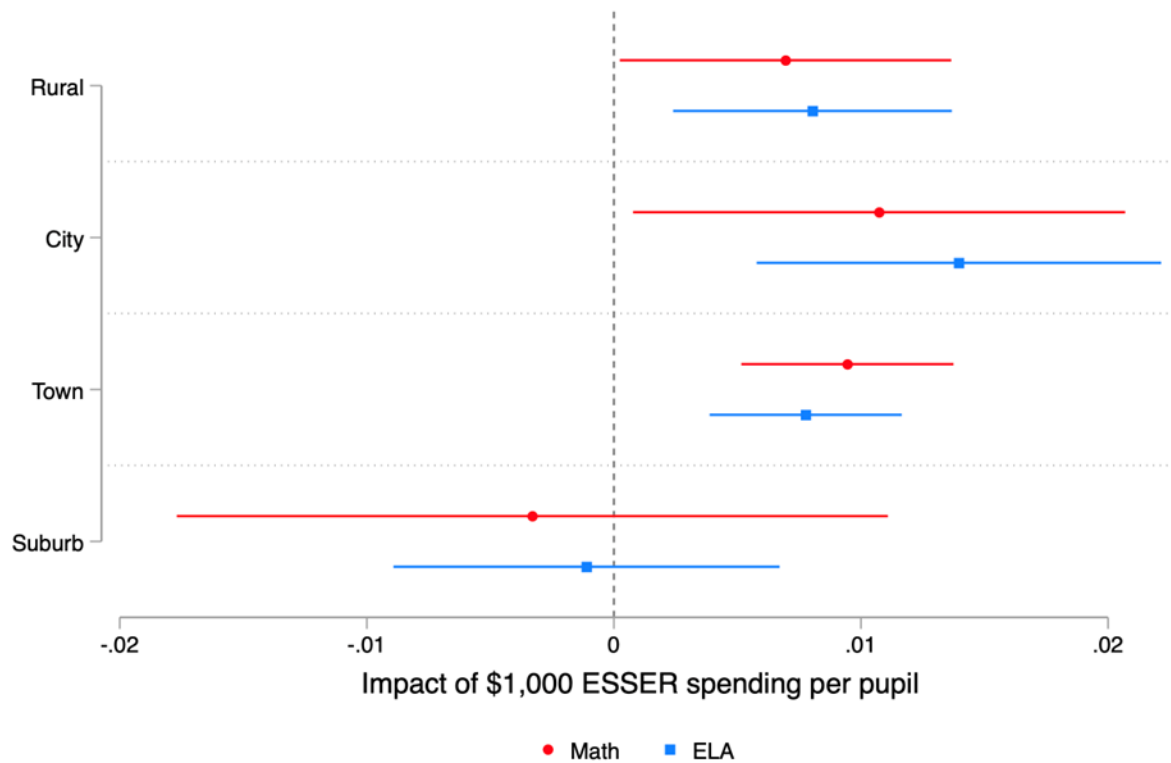
Note. Points represent estimates from our 2SLS model with state fixed effects, estimating on the subsample of districts grouped into quartiles according to 2019-20 district spending per pupil. Spikes present 95% confidence intervals

Figure A3. Estimates of heterogeneity across quartiles of pandemic learning loss



Note. Points represent estimates from our 2SLS model with state fixed effects, estimating on the subsample of districts grouped into quartiles according to the difference between standardized 2018-19 and 2021-22 test achievement in the indicated subject. Spikes present 95% confidence intervals.

Figure A4. Estimates of heterogeneity across quartiles of district urbanicity



Note. Points represent estimates from our 2SLS model with state fixed effects, estimating on the subsample of districts grouped according to Common Core of Data-reported urbanicity. Spikes present 95% confidence intervals.

Table A.1 States included in Stanford Education Data Archive achievement records and spending data

	Math sample	ELA sample	Spending data sample
Alabama	x		
Arkansas	x		x
California	x	x	x
Connecticut	x	x	x
Georgia	x	x	x
Illinois	x	x	x
Indiana	x	x	x
Kansas	x	x	x
Kentucky	x	x	x
Louisiana	x	x	x
Massachusetts	x	x	x
Michigan	x	x	x
Mississippi	x	x	x
Nevada	x	x	x
New Hampshire	x	x	x
New Jersey	x	x	x
North Carolina	x	x	x
Ohio	x	x	x
Oklahoma	x	x	
Oregon	x	x	x
Pennsylvania	x	x	x
Rhode Island	x	x	x
South Dakota	x	x	x
Tennessee	x	x	x
Utah	x	x	x
Virginia	x	x	x
Washington	x	x	x
West Virginia	x		x
Wisconsin	x	x	x
Wyoming		x	x

Note. State inclusion in sample demarcated with x's. Only states included in at least one of the samples listed.

Table A.2 Characteristics of districts for the average student by ESSER III allocation per pupil

	All districts by subject sample		Districts grouped by ESSER II & III allocation per pupil							
	Math sample	ELA sample	Math sample				ELA sample			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
ESSER II & III allocation per pupil (\$1,000s)	3.10	3.09	0.75	1.74	3.00	6.06	0.72	1.67	2.94	6.01
ESSER II & III spending per pupil (\$1000s)	1.17	1.17	0.28	0.62	1.05	2.38	0.26	0.61	1.03	2.36
Title I formula-eligible percent 2020-21	16.39	16.21	6.36	11.41	17.55	26.95	6.14	11.00	17.11	26.68
District enrollment (1000s)	37.76	38.11	15.29	27.35	38.18	62.85	14.51	27.95	38.80	62.46
% FRPL	48.47	48.41	22.82	41.66	57.16	66.09	22.18	40.11	56.72	66.86
% American Native	0.44	0.44	0.30	0.45	0.46	0.54	0.28	0.42	0.45	0.55
% Asian	6.15	6.26	11.01	6.25	4.86	3.35	11.34	6.26	5.18	3.35
% Black	14.26	14.35	5.39	9.30	13.37	25.73	5.24	9.37	13.07	25.84
% Hawaiian/Pacific Islander	0.32	0.31	0.25	0.29	0.55	0.21	0.24	0.28	0.50	0.21
% Hispanic	24.89	25.00	14.41	22.00	29.61	31.12	13.79	21.46	29.65	31.95
% Multiracial	5.21	5.22	5.70	5.64	5.42	4.30	5.68	5.72	5.38	4.39
% White	46.85	46.53	62.84	56.00	45.66	28.54	63.33	56.43	45.71	27.61
% Districts in rural area	15.41	15.03	11.84	16.35	20.94	12.64	12.25	15.21	20.57	12.24
% Districts in cities	30.41	30.00	11.52	22.28	31.21	50.68	10.77	20.81	30.10	50.90
% Districts in suburbs	43.53	44.59	71.39	49.08	34.90	25.00	72.12	52.07	36.52	25.53
% Districts in towns	10.64	10.37	5.26	12.29	12.95	11.69	4.86	11.91	12.82	11.34
District total revenue per pupil 2019-20 (\$1,000s)	15.54	15.63	16.09	14.93	14.43	16.53	16.26	14.99	14.50	16.59
Change in math 2019 – 2022 (SDs)	-0.15		-0.11	-0.14	-0.16	-0.18				
Change in math 2022 – 2023 (SDs)	0.05		0.04	0.04	0.04	0.06				
Change in ELA 2019 – 2022 (SDs)		-0.09					-0.07	-0.08	-0.10	-0.11
Change in ELA 2022 – 2023 (SDs)		0.02					0.03	0.01	0.01	0.03
<i>n</i>	4,989	4,824	1,248	1,247	1,247	1,247	1,206	1,206	1,206	1,206

Note. Means are weighted by district enrollment in 2022-23. All demographic and district characteristics are as of the 2022-23 school year. ELA=English language arts; ESSER=Elementary and Secondary School Emergency Relief Fund; FRPL=free/reduced-price lunch; SD=standard deviation.

Table A.3 Robustness of results to additional controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A. Math 2023 achievement											
ESSER per pupil spending	0.021***	0.020***	0.019***	0.016***	0.017***	0.019***	0.019***	0.018***	0.017***	0.018***	0.008*
2022-23 (\$1,000s)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
<i>n</i>	4,989	4,989	4,989	4,989	4,989	4,989	4,989	4,989	4,989	4,989	4,989
Panel B. English language arts 2023 achievement											
ESSER per pupil spending	0.018***	0.021***	0.020***	0.017***	0.013***	0.021***	0.018***	0.018***	0.024***	0.020***	0.003
2022-23 (\$1,000s)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
<i>n</i>	4,824	4,824	4,824	4,824	4,824	4,824	4,824	4,824	4,824	4,824	4,824
District % instruction remote and hybrid	X										X
State youth access to computer and internet		X									X
State COVID cases and deaths			X								X
State democratic vote share 2020				X							X
State unemployment rate					X						X
State total K12 education revenue 2019-20						X					X
State indicators for Right to Work							X				X
State additional Title I formula inputs								X			X
CRPE student performance saliency									X		X
State ESSER % spent prior to and during 2022-23										X	X

Note. Heteroskedasticity-robust standard errors are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. District covariates in all models include the following: district demographics for the 2022-23 school year including the portion who are American native, Asian, Black, Hispanic, Hawaiian or Pacific Islander, and multiracial; indicators for district urbanicity; district total revenue per pupil in 2019-20; and a cubic of achievement in the same subject as the indicated outcome from the prior school year. Model 1 includes separate controls for the share of instruction in 2020-21 conducted remotely and in hybrid format, omitting the share of the year in-person; data from the Return to Learn dashboard. Model 2 instead controls for the ACS 2018-22 share of youth with access to a computer at home and, separately, with access to internet at home. Model 3 controls for the state-level COVID cases and deaths in each school year starting in March 2020 through 2021-22 SY. Model 4 controls for the democratic candidate vote share at the state level in the 2020 presidential election. Model 4 controls for the state unemployment rate, averaged across each school year from 2019-20 through 2021-22. Model 6 controls for state total revenue from all sources going towards K12 education in 2019-20. Model 7 includes an indicator for Right to Work labor policies at the state level. Model 8 includes controls following Title I formula inputs: state per pupil expenditure on education; qualification for small state minimum provisions, by grant; and effort factor multiplier. Model 9 includes CRPE grades for student test performance saliency on state department of education websites. Finally, Model 10 includes state-level controls for the percent of all ESSER funding spent prior to 2022-23 and the percent spent in that school year. CRPE=Center on Reinventing Public Education; ESSER = Elementary and Secondary School Emergency Relief Fund; NCES = National Center for Education Statistics; SEDA= Stanford Education Data Archive.