Northwestern | INSTITUTE FOR POLICY RESEARCH

Working Paper Series

WP-18-02

Do School Spending Cuts Matter? Evidence from the Great Recession

C. Kirabo Jackson
Professor of Human Development and Social Policy
IPR Fellow
Northwestern University

Cora Wigger
Graduate Student
Northwestern University

Heyu Xiong
Graduate Student
Northwestern University

Version: January 9, 2018

DRAFT

Please do not quote or distribute without permission.

ABSTRACT

Audits of public school budgets routinely find evidence of waste. Also, recent evidence finds that when school budgets are strained, public schools can employ cost-saving measures with no ill-effect on students. The researchers theorize that if budget cuts induce schools to eliminate wasteful spending, the effects of spending cuts may be small (and even zero). To explore this empirically, they examine how student performance responded to school spending cuts induced by the Great Recession. They link nationally representative test score and survey data to school spending data and isolate variation in recessionary spending cuts that were unrelated to changes in economic conditions. Consistent with the theory, districts that faced large revenue cuts disproportionately reduced spending on non-core operations. However, they still reduced core operational spending to some extent. A 10 percent school spending cut reduced test scores by about 7.8 percent of a standard deviation. Moreover, a 10 percent spending reduction during all four high-school years was associated with 2.6 percentage points lower graduation rates. While the researchers' estimates are smaller than some in the literature, spending cuts do matter.

Do School Spending Cuts Matter? Evidence from the Great Recession*

C. Kirabo Jackson
Northwestern University and NBER

Cora Wigger
Northwestern University

Heyu Xiong
Northwestern University

January 9, 2018

Abstract

Audits of public school budgets routinely find evidence of waste. Also, recent evidence finds that when school budgets are strained, public schools can employ cost-saving measures with no ill-effect on students. We theorize that if budget cuts induce schools to eliminate wasteful spending, the effects of spending cuts may be small (and even zero). To explore this empirically, we examine how student performance responded to school spending cuts induced by the Great Recession. We link nationally representative test score and survey data to school spending data and isolate variation in recessionary spending cuts that were unrelated to changes in economic conditions. Consistent with the theory, districts that faced large revenue cuts disproportionately reduced spending on non-core operations. However, they still reduced core operational spending to some extent. A 10 percent school spending cut reduced test scores by about 7.8 percent of a standard deviation. Moreover, a 10 percent spending reduction during all four high-school years was associated with 2.6 percentage points lower graduation rates. While our estimates are smaller than some in the literature, spending cuts do matter.

I Introduction

For decades researchers have debated whether there is a causal relationship between increased school spending and students outcomes (Coleman, 1990). Until recently, much of the evidence on this had been correlational and suggested little association between school spending and student

*Jackson: Northwestern University, 2120 Campus Drive, Evanston IL 60208 (email:kirabo-jackson@northwestern.edu). Wigger: Northwestern University, 2120 Campus Drive, Evanston IL 60208 (email:CoraWigger2021@u.northwestern.edu). Xiong: Northwestern University, 2211 Campus Dr, (email:HeyuXiong2018@u.northwestern.edu). The statements made and views expressed are solely the responsibility of the authors. Wigger and Xiong are grateful for financial support from the US Department of Education, Institute of Education Sciences through its Multidisciplinary Program in Education Sciences (Grant Award # R305B140042).

outcomes (Hanushek 1997; Betts 1996). However, several well-identified studies now find that large permanent increases in school spending due to the passage of school finance reforms improved student longer run outcomes such as adult wages (Jackson et al., 2016) and crime (Johnson and Jackson, 2017), medium-run outcomes such as high school completion (Candelaria and Shores, 2017), and short run outcomes such as SAT scores (Card and Payne, 2002) and standardized test scores (Lafortune et al., 2018). This collection of studies is now being used both to to advocate for increased school spending, and also oppose reductions in schools budgets (Barnum, 2016).

However, audits of public school budgets routinely find evidence of waste and bloat (Ross 2015). This suggests that in the face of public school budget cuts, one *could* reduce school spending with little to no deleterious effects on children. The basic logic is that while schools may engage in wasteful spending when budgets are growing (e.g. Liebman and Mahoney 2017), hard financial times create incentives that induce firms and workers to curb wasteful spending and "*make do with less*" (Lazear et al. 2016). There is empirical support for this idea; in the wake of recessionary budget cuts, many districts induced expensive experienced teachers to retire early (Fitzpatrick and Lovenheim 2014), others deferred scheduled maintenance and eliminated non-essential travel (Ellerson 2010), and some districts operated schools for only four days per week (Anderson and Walker 2015).[1] Remarkably, both Fitzpatrick and Lovenheim (2014) and Anderson and Walker (2015) find that these cost-saving measures *improve* student outcomes; evidence that public schools may not always allocate their budgets in a manner that maximizes student achievement.

The increased use of cost-saving practices during recessions creates a puzzle. If such measures reduce costs with no ill-effects for students, why don't schools use them all the time? To rationalize these patterns, we theorize that if (a) there is waste and bloat in the system, and (b) schools are punished for poor outcomes more than they are rewarded for good outcomes, schools will be more likely to cut back on wasteful spending practices during recessions. A direct implication of this model is that even if spending increases lead to improved outcomes, spending reductions may have little or no deleterious effects. To test this notion empirically, we examine the causal effect of reductions in school spending due to budgetary cuts in response to the Great Recession.

Our main outcome of interest is student test scores from the restricted individual-level National Assessment of Educational Progress (NAEP) data. The NAEP, commonly referred to as "the Nation's Report Card," tests state-representative samples of over one hundred thousand students in math and reading in 4th and 8th grade every two years. We focus on exams given between 2002 and 2015. We link these test score data to district level-spending data from the Common Core of Data (CCD) between 2000 and 2015. To examine longer-run outcomes we also employ the Integrated Public Use Microdata Series (IPUMS) which contains information on high school completion among a representative sample of individuals in a state.

[1] As of 2016, this has been tried in over 100 districts across several U.S. states (Hill and Heyward, 2017).

To isolate plausibly exogenous variation in school spending, our identification strategy exploits the fact that states that funded high shares of their pre-recession school budgets from state sources were those that saw the largest reductions in per pupil spending in the wake of the Great Recession. That is, because state tax collections fell dramatically during the recession, states that were historically most reliant on state tax collections to fund public schools saw the deepest school spending cuts (Evans et al. 2017; Shores and Steinberg 2017). Our instrument is, therefore, the interaction between the share of state K12 spending that came from state sources in 2008 and the number of years post-recession. Our empirical strategy relies on the assumption that states with different levels of reliance on state revenue sources were not differentially affected by the recession for reasons other than through school spending. We present much evidence that our instrument is both relevant and likely exogenous. Specifically, we demonstrate that (a) both test scores and economic conditions in districts with high shares of state spending from state sources were on similar pre-recession trajectories as those with low shares, (b) high shares of state spending from state sources is a very strong predictor of reductions in school spending only *after* the recession, (c) changes in school spending as predicted by the instrument are unrelated to changes in unemployment or child poverty, and (d) adding controls for economic conditions at both the state and local levels has virtually no effect on our estimated relationships between school spending and student outcomes.

Using our instrumental variables strategy, we find that districts responded to spending cuts by disproportionately cutting more from construction expenditures and less from core K12 spending. While construction spending makes up only 5 percent of the average schools budget, it accounted for 47 percent of the reduced spending. Conversely, while current K12 operating spending accounts for 86 percent of all spending it accounted for only 55 percent of spending reductions. These patterns stand in contrast to those documented for spending increases due to school finance reforms (Jackson et al., 2016) and suggest that the marginal propensity to spend on different inputs varies when there are spending increases than when there are spending reductions. This is consistent with our model in which districts become more efficient at spending when times are hard. The results also suggest that the budget cuts were sufficiently large that, even if there were some bloat, districts were forced to make cuts that likely hindered student outcomes.

Looking at test scores, on average, a 10 percent reduction in per-pupil spending led to about 0.078 standard deviations lower test scores (p-value$<$0.01). Reassuringly, these effects are unchanged as one controls for the prevailing economic conditions; suggesting that our effects on test scores are driven by the changes in public school spending only. We test this further, by showing that while recessionary school spending cuts are strongly associated with reductions in *public* school test scores, they are not systematically related to changes in *private* school test scores. We also examine effects by grade level and subject. As is common in other settings, effects are larger for math than reading. A 10 percent reduction in per-pupil spending leads to roughly 3.55 percent of a standard deviation lower scores in reading and 11.83 percent of a standard deviation lower scores in math. These estimated effects are about half as large as Lafortune et al. (2018). [2] To put these effects into perspective, these effect sizes are roughly equivalent to that of reducing teacher quality by one standard deviation (i.e. having all teachers in the district go from average quality to the 15th percentile) (Jackson et al., 2014).

To examine effects on longer-run outcomes, we test for whether exposure to recessionary spending cuts had deleterious effects on high school completion using Census data. Within a difference in differences framework (comparing the difference in outcomes between exposed and unexposed cohorts in states that had smaller or larger recessionary budgets cuts), exposure to the larger recessionary school spending cuts reduced high school completion for those who were exposed during their elementary school years. More specifically, a 10 percent reduction in per pupil spending across all four of an individual's high-school-age years reduced the likelihood of completing high school by 2.7 percentage points. This is similar to estimates based on school finance reforms in recent years (e.g. Jackson et al. (2016) and Candelaria and Shores (2017)).

Overall, our theoretical framework and our analysis of school budgets suggest that the effect of spending cuts may be smaller than those of spending increases. At the same time, our analysis of student outcomes are inconsistent with spending cuts having no deleterious effect on student outcomes at all. Our findings contribute to the long-standing debate around whether school spending matters by showing that (a) spending cuts harm students, and (b) spending matters even when using variation other than that due to school finance reforms. A distinguishing feature of this work is that our results shed light on how school districts may react asymmetrically to increases versus decreases in budgets. Accordingly, these results contribute to the field of behavioral public finance. Finally, our results deepen understanding of the long-run effects of growing up during a recession. While it is well-documented that growing up during a recession can lead to long lasting ill-effects through channels such as parental job displacement (Oreopoulos et al. 2008; Ananat et al. 2011; Stevens and Schaller 2011) and increased food insecurity (Gundersen et al. 2011; Schanzenbach and Lauren 2017), we provide the first evidence that recessions have lasting ill-effects on young individuals through their effects on the governments' abilities to provide public education services.

The remainder of the paper is as follows. Section II present a theoretical framework within which the effects of spending reductions may differ from those of spending increases. Section III describes the data. Section IV describes the empirical strategy. Section V presents the main results, Section V.1 explores mechanisms, and Section V.2.2 presents robustness checks. Section VI presents some discussion of the findings and concludes.

[2] Lafortune et al. (2018) find that "The implied impact is between 0.12 and 0.24 standard deviations per $1,000 per pupil in annual spending." We take the midpoint of these values 0.18 as there overall estimate. Average spending levels in our period were $9540.35, so that a $1,000 is roughly a 10 percent increase.

II Theoretical Framework

We present a theoretical framework for thinking about why the marginal effect of spending increases could be large and positive, while those for spending reductions could be much smaller and possibly zero. Our framework highlights that for asymmetric spending effects to emerge requires *greater* incentives to reduce waste when budgets contract than when they are growing. Informed by the idea that voters may be loss averse (Kahneman and Tversky, 1979) and may pay more attention to bad outcomes than good (Lau, 1982), we present a framework with such asymmetric incentives.

In our simple model, a school administrator faces a reward system imposed by parents and chooses how to spend the district budget. The annual budget, B is exogenous. Districts spend τ on things that are valued by school administrators, but are not productive for students. We refer to this unproductive spending as "slack". It may take the form of job perks (such as nice coffee lounges, or non-essential travel) or inefficient spending (e.g. educators hold back on efforts to locate the cheapest textbooks or supplies). With the remaining budget, $(B-\tau)$, school districts maximize student output. To maximize output, districts spend the first productive dollar on the input with the highest marginal output per dollar spent, and then spend the next dollar on the the remaining input with the highest marginal output per dollar, and keep doing this until the productive budget is spent. Maximum student output (y^*) is therefore a concave increasing function of the productive budget such that $y^* = f(B-\tau)$, where $f' > 0$ and $f'' \leq 0$.[3]

School administrators value slack directly $g(\tau)$, where $g' > 0$, and receive some payoff based on the student outcomes from parents. Parents may be loss averse (Kahneman and Tversky, 1979) such that they punish test score reductions more harshly than they reward increases. This notion that public officials are differentially punished for bad outcomes has been documented by political scientists (e.g. Lau (1982) and others). For simplicity, we model this as a utility *penalty* (P) for outcomes that are worse than in the previous year. Where we index time with the subscript t, the resulting administrator's utility is as below.

$$\begin{aligned} U_t &= g(\tau_t) + f(B_t - \tau_t) \quad if \quad y_t \geq y_{t-1} \\ U_t &= g(\tau_t) + f(B_t - \tau_t) - P \quad if \quad y_t < y_{t-1} \end{aligned} \tag{1}$$

The school administrator's problem is to choose the level of slack τ to maximize their utility given the budget they receive and the reward structure imposed by parents.

Case 1: *Symmetric rewards (i.e. P=0)*

The problem simplifies to the case where the administrator chooses τ to equate the marginal benefit of one dollar in "slack" to the marginal payoff from the forgone outcome growth associated

[3] The assumption that $f' > 0$ is required for the the indirect outcome function to be invertible. This assumption is not required for the key result that spending increases may have a differently sized effect as spending reductions.

with that dollar of slack. That is, $dg/d\tau = -df/d\tau$. So long as τ is a normal good, the optimal slack level τ^* is an increasing function of B. As such, the the utility maximizing level of the outcome y^* is an increasing function of B. That is, $y^* = f(B - \tau^*(B)) = h(B)$. Under the assumption that y is a normal good $d[B - \tau^*(B)]/dB > 0$, so that $h' > 0$, and $h(B)$ is invertible. As such, the increase in equilibrium output y^* when the budget goes from B to $B+1$ is the same size as the reduction in output when the budget goes from $B+1$ to B. That is, with symmetric rewards, the effect of spending increases and decreases are opposite in sign and equal in magnitude.

Case 2: *Asymmetric rewards due to loss aversion (i.e. $P > 0$)*

From the assumption that both goods (τ and y_t) are normal, any spending increase (from B to $B+1$) will go toward increasing both test score growth and slack. Accordingly, for any budget increase, outcome growth is positive and (as in Case 1) the administrator chooses τ such that $dg/d\tau = -df/d\tau$. However, with the penalty P, the utility maximizing level of τ may be to have no reduction in test scores even in the face of a budget decrease. To see this, consider the following. As the budget falls from $B+1$ to B, the reduction in utility from cutting slack by the full 1 dollar is $dg/d\tau$, while the reduction in utility from spending 1 dollar less on test scores is $P + df/d\tau$. Generally, because $\lim_{d\tau\to 0}(df/d\tau) = 0$, for any budget cut δ, so long as the penalty (P) is larger than the overall loss in utility from cutting δ from slack, the administrator will avoid the penalty and will cut δ from slack. More formally, for any budget decrease from $B+\delta$ to B, with a sufficiently large penalty such that $P \geq \int_B^{B+\delta}[dg/d\tau]d\tau$, the administrator will reduce slack by the full amount of the cut to avoid the penalty, and will therefore leave test scores unchanged.

Note that where this condition does not hold (i.e. where the penalty is small or the budget cut is large), the administrator cuts spending on the margin as they would in the symmetric case. This implies that small budget cuts may have little effect on test scores, but that large cuts will have larger marginal effects. Also, given that $dg/d\tau$ is decreasing in τ, places with high levels of slack may have no test score response to spending reductions, while areas with low levels of slack may experience large test score responses. While these are important implications, our key insight is that the marginal effects of spending cuts may be much smaller than those of spending increases.

In sum, in a world in which (a) there is some slack in the budgets of schools, and (b) schools are punished for bad outcomes more than they are rewarded for good outcomes, the effect of spending increases need not be the same as that of spending reductions, and spending reductions may have no effect on student outcomes at all. Specifically, school districts may cut back on non-productive spending during times of budgetary contraction such that test scores do not fall. However, because most well-identified studies of the effect of school spending are based on spending increases, the extent to which this holds empirically is an open question. We test this notion empirically in Section V and present some evidence on mechanisms in Section V.1.

III Data

To examine the effect of school spending reductions on student test scores, we link individual-level achievement data matched to the Local Education Agency (LEA) and state finance data.[4]

Finance Data: School finance data come from the Annual Survey of School System Finances as reported by the U.S. Census Bureau. The surveys are conducted annually and aggregate financial revenue and spending data for public elementary and secondary schools at the Local Education Agency (LEA) and State Education Agency (SEA) level. LEA surveys include data from all public school districts (approximately 13,500), and state surveys include data from all 50 U.S. states and Washington D.C. The financial survey provides education revenue information broken down by local, state, and federal sources, along with overall student enrollments. These financial data are available between 1986-87 and 2014-15 and break down expenditures into broad categories, including instructional spending, capital expenditures, and payments to private and charter schools.

Student Achievement: Our measure of student achievement comes from the National Assessment of Educational Progress (NAEP) administered by the National Center for Education Statistics (NCES). The NAEP is referred to as the Nation's Report Card as it tests students across the country on the same assessments and has remained relatively stable over time. While state accountability assessments may vary across content and proficiency standards, the NAEP provides a measure of student achievement that can be compared across states, districts, and years. The NAEP is administered every other year to a population-weighted sample of schools and students.[5] We use restricted-use data files with individual-level NAEP scores. We link students' scores with their LEA and State, and are able to control for individual characteristics such as free/reduced lunch status and English learner status. We focus on public school students' 4th and 8th grade Math and Reading assessment scores, which are generally available every other year after 2000.[6] To facilitate comparisons over time, we report NAEP scores standardized to a base year of 2003. Because the NAEP sample has been increasing over time and only stabilized after 2000 (see Appendix Table 1), we focus on the period between 2002 and 2015. Our analytic student level sample consists of almost 4.3 million individual observations of NAEP scores linked to over 13,000 school districts in all 50 states and Washington, D.C. from the years 2002-2015 (see Table VI).

High School Completion Rates: While test score are our main outcome of interest, we supplement the test score data with high-school completion data from the Integrated Public Use Microdata Series (IPUMS). We collect data from the 2000 decennial census and the American Community

[4]We provide more extensive detail on our data sources in the Data Appendix.

[5]Schools are selected from 94 geographic areas, 22 of which are always the same major metropolitan areas. Students are selected randomly within the selected schools to complete the assessments. Note that our main results are invariant to the use of sampling weights.

[6]While we limit the sample to public school students only to correspond to our public education financial data, we do present a falsification using private school student test scores in Section V.2.2).

Surveys (ACS) from 2000 to the present. We compute the high-school completion rate for individuals at each age (between 18 and 30), in each state, in each survey year. The resulting dataset is at the age-state-year level, and summarizes the average high school completion rate in each cell.

Other Data: We supplement financial and achievement data with area and school-district demographic, employment, and economic characteristics. From the United States Census Bureau Small Area Income and Poverty Estimates (SAIPE) we obtain estimates on the total population, child population, and child population living in poverty for the geographic areas associated with school districts.[7] We also use area economic indicators of employment and wages from the Bureau of Labor Statistics (BLS). BLS reports data at the state and county level. To form district-level economic indicators we use weighted county-level estimates by the overlapping population within the county and corresponding school district. We also include public school district staffing information and student demographics from the Common Core of Data LEA Universe surveys from the National Center for Education Statistics (NCES).[8]

IV Empirical Strategy

Our goal is to estimate the causal effect of per pupil public school spending reductions on student achievement.[9] To this aim, we exploit plausibly exogenous variation in school finance within states over time that are induced by the Great Recession. The Great Recession began in December of 2007, and was the most severe economic downturn in the United States since the Great Depression. During the following 18 months, the unemployment rate increased by over 5 percentage points (Evans et al., 2017) and housing prices fell by about 30 percent (Kaplan et al. 2017; Hurd and Rohwedder 2010).[10] The sudden reduction in state revenues due to lower sales and income taxes led to cuts in state funding for education and other services (Business Cycle Dating Committee 2010; Chakrabarti et al. 2015; Leachman and Mai 2014).[11]

[7] SAIPE estimates are used for determining Title I eligibility, and are determined from federal tax data and either the American Community Survey (2005 and after) or the Current Population Survey (2004 and prior).

[8] When these data are missing, we impute with the mean. Reassuringly, imputing the mean for missing values does not affect the results in a meaningful way, but it does avoid arbitrarily dropping observations with missing data.

[9] The central challenge lies in the potential bias in the relationship between public school expenditures and student outcomes. Such biases may emerge for several reasons. In the cross section due to the local component of public school funding, (1) areas with higher wealth will tend to have higher levels of schools spending, and (2), conditional on local wealth, areas that value education and therefore have higher local taxes will spend more on education. To address this concern, most credible approaches rely on changes in school spending within geographical areas. However, such approaches may also be biased because changes in local demographics or changes in local economic conditions may mechanically lead to changes in school spending and also have a direct effect on outcomes. As such, the most credible approaches rely on variation in school spending within geography areas that occur for reasons largely external or exogenous to *other* changes within the geographic location itself. This is our approach.

[10] Over eight million private sector jobs were lost and employment did not return to pre-recession levels until 2014.

[11] Shores and Steinberg (2017) document that areas that experienced the largest reductions in employment during the recession also experienced greater test score reductions relative to areas that were less hit. Because the Great Recession had wide ranging effects in several domains (including parental employment, other services), this finding,

To isolate the impact of school spending reductions from other recessionary effects, we focus on changes in school spending caused by the Great Recession but were unrelated to broader recessionary effects in other domains. We are able to do this because the impact of the Great Recession on school district budgets differed from that on the broader economy in systematic ways. Evans et al. (2017) and Leachman et al. (2017) document that the structure of school finance systems prior to the recession moderated the severity of the recession on schools' finances. Districts that were heavily dependent on funds from *state* governments were particularly vulnerable to the recession.[12] To see why this is true, Figure 1 plots the national aggregate school revenue by funding source over time. While overall school spending declined during the recession (between 2007 and 2010), revenues from the major state taxes – the income tax and the sales tax – fell the most sharply during this period.[13] Because districts in different states relied on different proportions of money from local, federal, and state sources, states that were heavily reliant on state sources of revenue to fund public education tended to experience larger school spending reductions during the recession than those in states that relied more heavily on local funding or federal funding.

To see this pattern more clearly, the left panel of Figure 2 plots the state-level percent change in Per-Pupil spending between 2007 (pre-recession) and 2011 (post recession) against the share of K12 spending in the state that came from state sources. Following Evans et al. (2017), we compute the share of state K12 revenues in state s that came from state sources in 2008 (pre-recession) across all districts in the state as follows:

$$\Omega_s = \frac{\sum_{d \in s} State\ Revenue_d}{\sum_{d \in s} Total\ Revenue_d} \tag{2}$$

where $State\ Revenue_d$ denotes the school revenue in district d which came from state sources in the 2007-2008 school year; and $Total\ Revenue_d$ is the total revenue collected in district d in the same year. This variable captures cross-state variation in the reliance on state revenue to fund public K12 schools. Figure 2 shows a clear tendency for states that were more heavily reliant on state sources

while important, does not speak to the impact of school spending reductions *per se*. They do also find that areas that experienced the largest reductions in per-pupil spending during the recession tended to experience greater test score reductions. There is evidence that many states increased tax collection efforts in the wake of the recession (Ellerson, 2010) such that using the raw spending changes during the recession will be endogenous to decisions made by the district in response to recessionary revenue cuts. Accordingly, their estimated relationships may not be causal. Our approach does not use this potentially endogenous variation.

[12]Over the past 40 years, there has been a marked shift away from locally funded public schools toward state-financed public schools. This growing role of the states in education is in part an attempt to equalize education resources across districts. It is a response to a long series of court cases that have challenged the constitutionality of an education finance system that has led to wide disparities in education spending across school districts. For instance, Serrano I in 1971 and subsequent cases led to a requirement of equal spending per student in California.

[13]During this time period, owing to the American Recovery and Reinvestment Act of 2009 (ARRA), which sought to temporarily offset for the loss in state funding, education spending from federal sources increased in 2010 and 2011 and then fell back to pre-recession levels thereafter.

to have larger reductions in K12 spending during the recession.

This pattern motivates our instrumental variables approach. We use Ω_s as an exogenous shifter of K12 spending within states during the recession. For this approach to uncover a school spending effect, Ω_s should not be correlated with other policies or changes in economic conditions within states. We argue that the extent to which a state relied on state revenue sources to fund education prior to the recession is unrelated to the impact of the recession on *other* dimensions in that state. For the first test of this, Appendix Figure 1 shows that Ω_s is evenly distributed across the geographic regions of the United States. To asses this assumption more directly, the right panel of Figure 2 plots changes in the state unemployment rate between 2007 and 2011 by the share of K12 revenues in the state that came from state sources (i.e. Ω_s). In the average state, unemployment rates increased by about 3 points between 2007 and 2011. However, consistent with our contention, Ω_s is unrelated to the impact of the recession on *other* dimensions in that state. We also present additional tests of the validity of our instrument in Sections IV.2 and V.2.2.

IV.1 Estimation Equation

Exploiting this plausibly exogenous variation, we implement a within-state two-stage least squares (2SLS) regression model where we instrument for changes in per-pupil school expenditure during the post-recession years with the fraction of state education revenue from state sources during the pre-recession years, Ω_s. Intuitively, we compare the change in student achievement before and after the recession across states with a high or low fraction of revenue from state sources. If the only reason for a *change in* the difference in outcomes across areas with high and low Ω_s is the differential effect of the recession on public K12 spending, our instrument is valid. Formally, we estimate systems of equations of the following form by 2SLS.

$$\begin{aligned} PPE_{dst} = \sigma_1 \cdot (\ln(\Omega_s) \times I^{post} \times T) + \sigma_2 \cdot (\ln(\Omega_s) \times I^{post}) \\ + \sigma_3 \cdot (\ln(\Omega_s) \times T) + \pi C_{idst} + \gamma_d + \gamma_t + \varepsilon_{dst} \end{aligned} \tag{3}$$

$$Y_{idst} = \beta \cdot (PPE_{dst}) + \alpha \cdot (\ln(\Omega_s) \times T) + \Phi C_{idst} + \phi_d + \phi_t + \varepsilon_{idst} \tag{4}$$

The endogenous treatment, PPE_{dst}, is the log per-pupil school spending in district d in state s during year t. The outcome Y_{idst} is the test scores for student i from district d of state s in year t. Using only variation across cohorts within districts we include district fixed effects γ_d and ϕ_d in the first and second stage, respectively. To account for general differences across years, we include year fixed effects γ_t and ϕ_t in the first and second stage, respectively. T is a scalar in the calendar year, and I^{post} is a post-recession indicator denoting all years after 2008. The excluded instruments in the first-stage are the interactions between $ln(\Omega_s)$ and the post recession timing variables I^{post} and $T \times I^{post}$. To account for any pre-recession time trend differences between high and low Ω_s

states, we include $ln(\Omega_s) \times T$ as a control. Note that our results are very similar across models that control for pre-trending and those that do not. This suggests that the assumption of parallel trends is likely satisfied on our data, and supports a causal interpretation of our results.

We also include vector C_{idst}, which is a set of district level socio-economic and demographic characteristics. These include the size of the residential population of the district, and the number of school-age children residing in the district (*this is not student enrollment*). Also, to account for underlying economic conditions, this includes differential time-fixed effects by the state's unemployment rate in 2007 (prior to the recession) and differential time-fixed effects by the districts' percent of school age children living in poverty in 2007 (prior to the recession). To account for possible direct effects of the recession itself, we follow Yagan (2017) and create a Bartik predictor.[14] Finally, C_{idst} also includes a set of individual-level controls. These include Limited English Proficiency (LEP) status, and Free Lunch status (a rough proxy for parental income).

Our linear model is motivated by general patterns in the data. To show how school finances responded to the Great Recession by pre-recession state revenue, we visually present a flexible DiD event study. Where X_{dst} is the subset of the district and county level controls from C_{idst}, we estimate the model below by OLS on our district-level panel.

$$PPE_{dst} = \sum_{t=2000}^{2015} \beta_t \cdot (\ln(\Omega_s) \times I_{T=t}) + \Pi X_{dst} + \gamma_d + \gamma_t + \varepsilon_{dst} \tag{5}$$

In equation 5, $I_{T=t}$ is an indicator denoting if the observation is for calendar year t so that the coefficients β_t map out the differences in public school spending between states with high and low Ω_s in each year. We plot these coefficients in Figure 3 where the reference year is 2007. The results indicate that in years prior to the recession, school spending was on a similar trajectory in areas with different levels of Ω_s, but that following the Great Recession, districts in states with heavy reliance on state revenue experienced a clear drop in per-pupil spending. This is a visual representation of our first stage. We present the formal first-stage regression results for our NAEP sample below.

IV.2 First Stage Regressions

Testing Instrument Relevance: In Table 2, we examine the first-stage relationship in the NAEP student sample. We estimate the model with per pupil spending in both logs and levels. Column 1 of Table 2 shows the coefficient on the interactions between Ω_s and the post-recession timing variables in predicting the level of K12 per pupil spending. The first notable pattern is that the overall Ω_s time trend is not statistically significant. This indicates that prior to the recession, states

[14] We compute the proportion of all workers in each industry in the county prior to the recession in 2007. Then for each industry, we multiply the industry proportions (defined at the county level based on 2007 values) with the national unemployment rate in that industry for each year. We then sum these products across all industries in each year to obtain our Bartik Predictor. See the Data Appendix for further detail.

with high and low levels of Ω_s were on a very similar trajectory of per pupil spending. However, the statistically significant coefficients on the interactions between Ω_s and the post-recession timing variables reveal that after the recession states with high and low levels of Ω_s were on very different trajectories of per-pupil-spending (specifically, areas with high reliance on state revenues experienced larger reduction in per pupil spending). The first stage F-statistic for this model is 34.55. Column 3 presents this same model in predicting the log of per-pupil-spending. The point estimates reveal the same pattern, and the first stage F-statistic is 25.29. It is worth noting that in the log specification, there is a small but significant pre-trend. While our main results are very similar with and without the inclusion of the linear trend, this suggests that (to be conservative) our preferred models should include a linear trend in the share of revenue from state sources in 2008.

Hawaii has only one district within the state (the Hawaii Department of Education) and it is an outlier state regarding the level of state spending (in fact, it is excluded from many studies of school spending). Given this outlier status, we estimate the first stage except that we interact the instruments with a ”Not Hawaii” indicator - Columns 2 and 4. The interactions reveal that the instruments predict spending changes in all states but *much* more so in Hawaii. In models that interact our instruments with ”Not Hawaii” indicators to capture this variation, our first stage F-statistics are well over 100. To assuage any concerns that our results are driven by the interactions with the ”Not Hawaii” indicators, we verified that we obtain similar results in models that (a) do and do not interact our instrument with Hawaii indicators, (b) models that both include and exclude Hawaii from the sample entirely. As such, our chosen specification uses all the available data and interacts the instruments with ”Not Hawaii” indicators (i.e. we use the model that produces similar point estimates as the simple model, but produces a much stronger first stage).

Testing Instrument Exogeneity: Table 2 shows the relevance of our instruments for changes in school spending in the post-recession years. However, the interpretation of β as the causal effect of school spending in the second stage also requires that the exclusion restriction holds. The credibility of our research design hinges on the assumption that states with different dependence on state sources were not differentially affected by the recession for reasons other than through school spending. We do provide some evidence of this. Specifically, we test whether instrumented school spending predicts the state unemployment rate, county child poverty (logs and levels), and county employment (logs and levels) in Table 3. We find no patterns of changing economic conditions in localities associated with our recession-induced changes in school spending. For all outcomes, the coefficients on log per-pupil spending (our preferred specification for predicting test scores) are never significant at the 5 percent level, and there is little evidence of differential time trending. This suggests that the size of the predicted spending changes is unrelated to either differential pre-trends or changes in the trajectory of economic outcomes. This lends credibility to our research design. We present additional tests in Section V.2.2.

V Results

V.1 Mechanisms

Before examining test scores, we test some of the implications of the model regarding how districts responded to budget cuts. To this aim, we use our 2SLS specification from equations 3 and 4 to estimate the extent to which different spending and staffing categories were reduced in response to recession-induced expenditure decreases. Table 4 and Table 5 report the results of separate 2SLS models estimated on all districts for which data is available, weighted by school enrollment.

Table 4 demonstrates how districts differentially allocate budget cuts across categories of spending. To show this, we regress the level of spending in each sub category (in per-pupil units) on the overall (instrumented) level of spending (in per-pupil units). The resulting coefficient is therefore the marginal propensity to spend in each category for each dollar increase in the overall budget. A convenient feature of this specification is that it allows for the formal test of whether the marginal propensity to spend in any category is equal to the average propensity to spend in that category. Insofar as the marginal and average propensities are different, it *may* suggest that districts (on the margin) respond differently to spending increases than they do to spending reductions.

Columns 1-3 show the categories that, combined, make up Total Expenditures per pupil. For every dollar in per-pupil spending cuts, districts decrease their capital expenditures by \$0.449 and their elementary/secondary expenditures by \$0.554. While these cuts suggest similarly dividing cuts across capital and elementary/secondary spending, they are highly disproportional to the average shares of these categories for overall expenditures. While capital outlay expenditures account for approximately 7.4% of overall per-pupil expenditures, they make up 44.9% of the allocation of reduced spending, suggesting that districts cut capital spending more than other forms of spending on the margin. The difference between the average share of capital spending and the share of capital spending that is cut when budgets are reduced is statistically significant with a p-value of 0.017. While this is *suggestive* of asymmetric spending effects, it is not dispositive because the marginal and average propensities to spend in each category may differ for reasons other than the asymmetric effects predicted by the model. In the ideal, one would have both spending increases and spending reductions and then one could compare the marginal propensities to spend for increases to that of decreases. To do this, we look at existing studies. By way of contrast, Jackson et al. (2016) find that when a school district received increased revenue due to a school finance reform each dollar increase in total spending was associated with \$0.1 increased spending on capital (a marginal propensity very similar to the average). This provides compelling evidence that when faced with a spending cut (as opposed to a spending increase) school districts are much more likely to cut from capital on the margin. To examine this result further, we estimated effects on construction and non-construction capital spending (columns 4 and 5). All of the reduction in capital spending was from construction. The disproportionate cutting of construction projects is consistent with the descriptive patterns documented in Leachman et al. (2016) and reports in the popular press that budget shortfalls forced school systems to defer maintenance and new construction.

This large reduction in construction spending is notable, because it suggests that school districts were able to preserve more of their core operational services by delaying construction projects. For a district that has a 10 percent budget cut (about \$1000 per-pupil for the average district), by cutting \$449 of that from construction, they are able to cut spending on core operations by only \$554 (only about 6.6 percent of core elementary/secondary spending). Consistent with this, elementary and secondary current spending accounts for 86.4% of overall per-pupil expenditures, but only 55.4% of spending cuts. This difference between the average and the marginal propensities to spend is statistically significant with a *p*-value of 0.035.

To gain a more detailed sense of what services were cut, columns 6 to 11 show the allocation of spending cuts across additional sub-categories of expenditures. For every dollar in exogenous spending cuts, districts reduced instructional spending by \$0.439 on average. Looking within instruction spending categories (columns 9 to 11) roughly half of this reduction can be accounted for by a reduction in instructional salaries. Notably, this marginal reduction in instructional salaries is real, but is less than the average share. Consistent with asymmetric spending effects, we can reject that the marginal decrease is the same as the average share with a *p*-value of 0.032. The rest of the reduction in instructional spending is accounted for by reductions in benefits and other types of instructional spending. In contrast to the reduction in instructional spending, while support services account for about 30 percent of spending on average, spending in this category was unchanged. The difference between the average share and the marginal share is statistically significant with a *p*-value of 0.038. These results stand in stark contrast to previous research on the allocation of exogenous spending increases resulting from school finance reforms. Jackson et al. (2016) demonstrate that funding shocks that increased spending resulted in slightly disproportionately higher increases to instructional spending and support services, whereas our results suggest that funding shocks that decrease spending result in proportional cuts to instructional spending and disproportionately lower, if any, cuts to support services.

While the reductions in instructional salaries suggests that student outcomes may suffer, this need not be the case. Given that reductions in instructional salaries and benefits could have been due to the hiring of fewer staff (which would likely affect outcomes) or the hiring of cheaper staff (which could have little effect on outcomes), it is important to look at staffing directly. Table 5 shows the 2SLS estimates of log per-pupil spending on staffing per pupil. When spending is reduced by 10 percent, teachers per-pupil are reduced by approximately 0.04 (a decrease of about 5.8 percent), while other staff categories are unchanged. These findings are consistent with significant decreases in instructional salaries (and no change in support service spending) documented in Table 4. Simple calculations based on Table 4 reveal that a 10% reduction in per-pupil spending led to a 6 percent percent reduction in instructional spending - indicating that the reductions in spending were due to hiring fewer teachers (rather than hiring less expensive teachers). Our finding that the biggest decreases in staffing occurred through reducing the number of teachers per-pupil is consistent with other evidence finding that the number of employed public school teachers decreased after the recession (Evans et al., 2017)

The patterns suggest that in the face of a budget cut, districts first reduced spending on construction projects, cut the remainder from instructional spending (largely by reducing instructional salaries and benefits through teacher staffing cuts), and left spending on support services relatively untouched. The lack of a reduction in spending on support staff is surprising. We speculate that support staff may have been necessary to allow districts to cope with the cuts in other areas. Compared to research on how funding increases are distributed across categories of expenditures, these results suggest that districts may allocate their budget changes differentially depending on whether the change is in cuts or increases. Insofar as construction projects yield little educational benefit (at least in the short run), this pattern is consistent with the predictions of the model such that school districts cut back on non-essential spending first before having to cut core services that may have deleterious effects on student outcomes.

V.2 Test Scores

Before turning to the parametric 2SLS results, we first estimate flexible models to provide visual evidence of basic patterns in our data. Figure 4 presents the event-study estimates of the reduced-form effects of our instrument. Specifically, we estimate the model below by OLS.

$$Y_{idst} = \sum_{t=2002}^{2015} \beta_t \cdot (\ln(\Omega_s) \times I_{T=t}) + \Pi C_{idst} + \gamma_d + \gamma_t + \varepsilon_{idst} \tag{6}$$

As with equation (5), the coefficients β_t map out the differences in test scores between states with high and low Ω_s in each NAEP testing year between 2002 and 2015. We plot these coefficients in Figure 4 where the reference year is 2007 (the testing year prior to the onset of the recession).

Overall, a clear pattern emerges. Student test scores in states with high dependence on state revenues (revenue collected primarily from state income and sales taxes) to fund public K12 schools declined following the recession, relative to areas that were less reliant on state revenues to fund public K12 schools. While the semi-parametric individual year effect interactions are somewhat noisy, the linear fit confirms the existence of a trend break. Reassuringly, during the pre-recession years, there is no discernible differential trending in test scores by instrument dosage. This indicates that student outcomes in states that relied on revenues raised from primarily state sources were on a very similar trajectory as other districts until the onset of the recession. However, in states with greater reliance on state revenues for public school funding (and therefore saw greater declines in per-pupil school spending), student performance dropped following 2008, the start of the recession, and continued to decline thereafter.

Having established visually the decline in post-recession outcomes for districts in states with high reliance on state revenues, we now use this relationship to quantify the causal relationship between student spending and student performance. The resulting 2SLS instrumental variables models provide a direct estimate of the causal impact of school spending and allow for formal tests of statistical significance. Table 6 presents the 2SLS estimated effects of spending on NAEP test scores. The dependent variable across all specifications is individual student test scores, standardized to 2003 scores. We show the marginal effects of spending in both levels and logs.

Column 1 of Table 6 presents the effect of per-pupil spending in levels (thousands of dollars). This is the most parsimonious model with no linear trend in state dependence on state revenues, no controls for economic conditions, and no Bartik predictor. In this parsimonious model, the coefficient of 0.06 indicates that for every \$1,000 decrease in per student spending, test scores declined by 6 percent of a standard deviation (p-value$<$0.01). Columns 2 through 4 progressively add more controls to illustrate the robustness of our result. Columns 2 adds the predictors of economic conditions and Bartik predictors, and column 3 adds a linear pre-trend for the instrument. The point estimate is largely unchanged with the addition of these controls. It is worth noting that the linear trend itself is not statistically significant. Given the lack of a visible pre-trend in Figures 3 and 4, this is unsurprising. However, it does bolster our claim that our variation is exogenous. To assuage any lingering concerns that our estimates are confounded by underlying recession intensity, column 4 present results that control for economic conditions directly. Note that we consider this model ”over controlling”, but present it only to establish the robustness of our result. This model includes the state unemployment rate, the county employment level, and the level of child poverty in the district. As one can see, the point estimate is virtually identical to our preferred specification in column 3 with the inclusion of these variables. This is consistent with the patterns in Figure 2 such that conditional on exposure to the recession (accounted for with year fixed effects), our instruments are unrelated to recession intensity itself.

Due to diminishing marginal returns to school spending, the log of spending is often a better predictor of outcomes than levels (Jackson et al., 2016). Column (5)-(8) show the effect of log spending on standardized student test scores. In the parsimonious model with few controls, the coefficient of 0.91 indicates that for every 10 percent decrease in per student spending, test scores declined by 9.1 percent of a standard deviation (p-value$<$0.01). Adding predictors for economic conditions has no appreciable effect on this estimate (column 6). Including the linear trend reduces the coefficient to 0.78, but the effect is very similar (column 7). In this conservative but preferred model, the coefficient of 0.78 indicates that for every 10 percent decrease in per student spending, test scores declined by 7.8 percent of a standard deviation (p-value$<$0.01). In the ”over-controlling” model that also includes the economic variables, the coefficient is statistically significant at the 1 percent level and is virtually identical to our preferred model without the economic variables.

Per-pupil spending was about \$10,000 on average during our sample period. As such, our log spending estimates suggest that decreasing school spending by \$1000 (about a 10 percent change) would reduce test scores by roughly 7.8 percent of a standard deviation. This is very similar to the results from the linear model. The facts that (a) we obtain the same basic result in linear models and in linear-log models, (b) we have no pre-trending in our outcomes for more and less treated states, (c) our instruments do not predict economic variables, and (d) our estimated effects are robust to the inclusion of a rich set of economic controls, suggest that this is a robust finding and that our estimated effects can be interpreted causally. We present one final test in Section V.2.2.

V.2.1 Effects by Grade and Subject

To explore heterogeneity, we estimate effects separately by grade and subject. We focus on our more conservative preferred model that includes linear pre-trends in our instrument and the predictors of recession intensity (we do not include the economic variables directly as covariates such as these may be over-controlling). Table 7 presents the 2SLS/IV coefficients estimated separately by subject, grade, and subject-grade. Each column corresponds to Column (7) of Table 6 restricted to the sample specified. The effect of per-pupil spending is statistically significant for each sub-sample of the data. However, the effect of spending is statistically significantly greater in Math as compared to Reading. The point estimates suggest that decreasing school spending by 10 percent would reduce math test scores by 11.8 percent of a standard deviation but reading scores by only 3.5 percent of a standard deviation. These results are consistent with previous studies which have shown that math scores are more elastic to spending and evidence from early childhood interventions. To put these effects into perspective, these effect sizes are roughly equivalent to that of reducing teacher quality by one standard deviation (i.e. having all teachers in the district go from average quality to the 15th percentile) (Jackson et al., 2014).

We also examine effects by subject and grade. The largest effects are observed for 8th and 4th grade math, where the effects coefficients are 1.455 and 1, respectively. This indicates that reducing public school spending by 10 percent reduces math test scores by 14 and 10 percent of a standard deviation in 8th and 4th grade, respectively. The effect is larger for 8th Grade performance relative to 4th Grade, although not statistically significantly so. The smallest effects are observed for 4th and 8th grade reading. The point estimate indicates that reducing public school spending by 10 percent reduces reading test scores by 2.92 and 4.65 percent of a standard deviation in 4th and 8th grade, respectively (both significant at the 1 percent level).

V.2.2 Robustness checks

Sorting bias: One worry in papers that analyze school spending at the district level with aggregate data is that the results could be biased by selective migration. For example Lafortune et al. (2018) state that they *"cannot rule out small effects of SFRs on student sorting"* in their analysis and Candelaria and Shores (2017) note that the district level CCD data *"may not reveal sorting within districts."* Our results are robust to any sorting across districts because our treatment occurs at the state level and impact all districts in the state. Unlike these other studies, we are able to rule out any selective migration across districts within states by aggregating the data to the state level and seeing if the effects persist for the entire state.

An additional bias not addressed in existing studies is sorting across sectors (i.e students moving from the public to the private school sector or *vice versa*). We can address this second type of sorting because the NAEP data also collect information from students at private schools. As such, we can aggregate both the public and private school data to the state year level and analyze both together. In principle, if stronger students from the public school sector were more likely to exit public schools for private schools at the same time that public schools lost funding, it could generate the patterns we document. However, if this were driving the results, then there would be no effect on aggregate test scores (both public and private) for the state as a whole. We present our 2SLS analyses on the combined public and private state-level data in Table 8.

Column 1 of Table 8 shows the effect of both public and private schools combined excluding the linear pre-trend in the instrument. The coefficient for the combined data is positive, statistically significantly different from zero, and very similar to the effect on public school scores in Table 6. Column 2 adds controls for economic conditions and the Bartik instrument, while column 3 additionally controls for the linear trend, analogous to column 7 of Table 6. In all models the coefficient for the combined data is positive, statistically significantly different from zero with a p-value of 0.1, and similar to the effect on public school scores in Table 6. The addition of economic controls in column 4 leaves these estimates largely unchanged. These results provide evidence that our overall test-score results in Table 6 are not a result of sorting within states.

Falsification test: The private school data can also be useful in testing the basic mechanisms behind our effect. Even though we show that our instrument is not correlated with changes in economic conditions and we show that our estimates are robust to including controls for economic conditions, one may still worry that there are other changes driving our estimates. To test for this, Column 2 breaks out the public school and private test scores separately. In principle, if our effects operate through reduction in public K12 spending, we should observe test score effects for public schools and no effect for private schools. Columns (5)-(8) of Table 8 present the estimates for both private and public schools and follow the same specifications of control variables as columns (1)-(4). Column 5 excludes the linear time trend for the instrument as well as economic conditions.

The coefficient for spending on public school scores is 0.796 (p-value$<$0.05) while that for private school scores is 0.187 and statistically insignificant. A test of equality of effect across the two sectors yield a p-value of 0.04. Column 7 presents the analogous result including the linear trends for the instrument and predictors of recession intensity. The pattern of results are similar. In this model, a formal test of equality of effect across the two sectors yield a p-value of 0.038. In sum, though the point estimate for private school scores is positive, one cannot reject that the effects is zero, and one can reject that the effect on public schools is the same as that for private schools at the 4 percent level. This is compelling evidence that our effects are indeed driven by changes in public school spending and not other potentially confounding factors.

V.3 High School Completion

While test scores have often been the focus of school finance studies, decreased test scores may not translate into longer-run effects.[15] To assess this, we use Census data from IPUMS to examine the extent to which education spending cuts caused by the Great Recession affect high school graduation rates. Because high school completion likely reflects the cumulative effect of several years of educational inputs, we move away from the contemporaneous model and analyze the graduation rather of individuals who were exposed to the recession for different amounts of time. We leverage the fact that, in a sample of the full population, at the same point in time individuals from different birth cohorts within the same state varied in the number of school-age years (i.e. years between the ages of 6 and 18) that were exposed to the recession. If our test score effects are real, cohorts with more years of exposure should have lower graduation rates than those with fewer years of exposure, and the exposure effect should be largest in those states that were the most dependent on state revenues to fund public schools.

To present visual evidence, we implement an event study based on the difference-in-differences variation to estimate the effect of recessionary spending cuts on high school attainment. More formally we estimate the following model by Ordinary Least Squares (OLS):

$$Y_{sct} = \sum_{e=-5}^{7} \beta_e \, I_{E_c=e} \cdot \ln(\Omega_s) + \lambda_s + \lambda_c + \lambda_t + \varepsilon_{sct} \tag{7}$$

In (7), Y_{sct} is the high school completion rate of individuals from state s from birth cohort c in calendar year t. E_c is our measure of exposure and is the year an individual from birth cohort c would have turned 18 (i.e. the last expected year of high school) minus 2008 (the year of the onset of the recession). As such, for individuals born in 1990 (with expected graduation in 2008), E_c is 0, for those born in 1985 (with expected graduation in 2003), E_c is -5, and for those born in

[15] Indeed, Jackson (2018) finds that teacher impacts on test scores are much weaker predictors of their impacts on high school completion than teacher's impacts on grades, attendance, or discipline.

1997 (with expected graduation in 2015), E_c is +7. Because we are interested in how exposure to the recession varied among those exposed to larger or smaller recession induced spending cuts, we interact this measure of exposure with our measure of a state's reliance on state revenues to fund education ($ln(\Omega_s)$). Specifically, we include the interaction between ($ln(\Omega_s)$) and $I_{E_c=e}$, which is an indicator denoting each individual value of our exposure variable. To account for differences across states, any time effects that may affect the outcome, and the direct effect of exposure, we include state fixed effects (λ_s), year fixed effects (λ_t), and birth cohort year fixed effects (λ_c).

Because the birth cohort year fixed effects control for years of exposure directly, the coefficients on the interactions, the different values of β_e map out the *difference* in outcomes among cohorts with the same temporal exposure to the recession but who faced different recession induced school spending cuts. If there are no differential trends in graduation rates in the high versus low Ω_s states for the pre-exposure cohorts (i.e. cohorts that were 18 years old or older in 2008 with $E_c < 0$), the interactions should be the same for all the cohorts with no exposure to the recession. Also, if the reduced spending associated with recessionary budget cuts had a deleterious effect on graduation rates, then the interactions should become increasingly negative for cohorts with greater exposure to the recession.[16] To show such patterns visually, we plot the β_e on graduation rates by years of school-age exposure for birth cohorts that were expected to graduate high school before during and after the recession in Figure 5. Consistent with a causal effect, and consistent with the test scores results, the difference in graduation rates between high and low Ω_s states is stable among cohorts that would have graduated high school prior to the recession, while exposed cohorts in high Ω_s states experienced differentially lower high-school graduation rates.

Using this variation in a 2SLS framework, we quantify the marginal effect of a change in per-pupil spending on graduation rates. Because high school completion reflects spending across an individual school-age years, we use two cumulative measures of per-pupil spending. The first is the total per-pupil spending in an individuals school district during the ages 14 to 18 (i.e. spending during the high-school years), and the other is the total per-pupil spending in an individuals school district during the ages 6 to 18 (i.e. spending during all school-age years). Using both measures, we estimate the following system of equations by 2SLS.

$$\begin{aligned} ln(PPE_{sc}) = \sigma_1 \cdot (\ln(\Omega_s) \times I_{(E_c>0)} \times E_c) + \sigma_2 \cdot (\ln(\Omega_s) \times I_{(E_c>0)}) \\ + \sigma_3 \cdot (\ln(\Omega_s) \times E_c) + \pi C_{sct} + \gamma_s + \gamma_c + \gamma_t + \upsilon_{sct} \end{aligned} \quad (8)$$

$$Y_{sct} = \beta \cdot (PPE_{sc}) + \alpha \cdot (\ln(\Omega_s) \times E_c) + \Phi C_{sct} + \phi_s + \phi_c + \phi_t + \varepsilon_{sct} \quad (9)$$

[16] To show a first stage visually, we estimate this model on per-pupil spending between the ages 15 and 18 (the last four years of expected schooling). We plot the β_e on the log of per pupil spending by years of school-age exposure in Appendix Figure 3. The difference in per-pupil spending between high and low Ω_s is stable prior to the recession, but after the recession, exposed cohorts in high Ω_s states experience lower levels of per-pupil spending.

The endogenous treatment, $ln(PPE_{sc})$, is the log average per-pupil school spending in state s for birth cohort c. This is either average spending during the expected high school years or average spending across all 12 expected school-age years. The outcome Y_{sct} is the percentage of high school graduates for birth cohort c from state s in year t. Using only variation across birth cohorts within states we include state fixed effects γ_s and ϕ_s in the first and second stage, respectively. To account for general differences across years, we include year fixed effects γ_t and ϕ_t in the first and second stage, respectively. C_{sct} is a vector of state level socio-economic and demographic characteristics.

Similar to our test score models, our excluded instruments are measures of exposure to the recession (E_c and $I_{(E_c>0)}$) each interacted with our measure of treatment intensity Ω_s. E_c is our continuous exposure variable as defined above, and $I_{(E_c>0)}$ is an indicator denoting cohorts that have any exposure to the recession prior to the age 18. To account for underlying trend differences in outcomes between high and low treatment states we also include a linear time trend in percent state ($\ln(\Omega_s) \times E_c$) as a control. Our model is therefore identified off a change in the linear trend after the recession across states with high and low dependence on state funds in 2008. This is simply a linearizion of the patterns depicted in Figure 5.

The 2SLS results are presented in Table 9. We first use per pupil spending during the last four years of expected school as our treatment. In column 1, we exclude a linear trend for exposure (note that the event study figures indicate no pre-trend) and other recession-intensity predictors and obtain a coefficient of 0.242 (p-value$<$0.05). This suggests that reducing per-pupil spending by ten percent across the last four years of schooling reduces high school graduation by about 2.4 percentage points. In column 2 we add predictors of recession intensity and in column 3 we add a linear cohort trend in $ln(\Omega_s)$. As expected, these controls have little effect on the coefficient. In column 4 we also include controls for the unemployment rate, the annual average employment, and the child poverty rate in the state; this has virtually no impact on the estimated relationships. In this most conservative model, reducing spending by ten percent across the last four years of schooling reduces the high school graduation rate by about 2.6 percentage points (p-value$<$0.01).

Based on Figure 5, the effects of the recession are most apparent among those who were in elementary school at the time of the recession. As such, we also implement models where the treatment is the level of per pupil spending across all school age years (ages 6 through 18). As expected, in such models the estimates are larger. With no linear trend in exposure or other predictors of recession intensity, cutting spending across all school-age years by 10 percent reduced graduation rates by 4.7 percentage points (p-value$<$0.01). The addition of controls for predicting recession intensity (employment and child poverty in 2007 and Bartik predictors) changes this estimate only slightly. With the addition of a linear trend in the instrument, this estimate grows to 0.716 (p-value$<$0.1) but is less precise. With the economic variables, the estimate is very similar. In sum, cohorts that were exposed to greater recession-induced spending cuts while they were in school graduated high school at lower rates. This is consistent with the patterns for test scores.

VI Discussion and Conclusions

The policy and scholarly debates regarding whether public school spending matters have been going on for decades. However, using large permanent increases in public school spending caused by school finance reforms, recent papers have uncovered credible evidence of a causal link between increases school spending and improved student outcomes (Jackson et al. 2016; Candelaria and Shores 2017; Lafortune et al. 2018; Card and Payne 2002). Other recent studies exploit the fact that school budgets (through school finance formulas) are often driven by random variation in housing prices, inflation, or student enrollment to isolate plausibly exogenous variation in school spending (Hyman 2017; Miller 2017; Gigliotti and Gigliotti 2017). These studies also find that increased school spending improves student outcomes.

Despite this growing consensus, there has been no study on how school districts respond to large persistent cuts to spending and how these responses translate into student outcomes. Recent studies document that in response to budget cuts, school districts employ cost saving strategies that have no ill-effects on students. Motivated by this, we present a theoretical framework in which large persistent spending cuts may have small effects on student outcomes, even if a similarly sized increase would improve outcomes. We then test this notion empirically. Using recession-induced public school spending cuts we find that (a) consistent with our model - school districts do respond to budget cuts by cutting non-core operations spending first, (b) school districts were able to minimize cuts to core operating expenditures, but did have to cut such spending, and (c) student performance was hurt by recessionary budget cuts.

Our results present further evidence that there is a causal link between changes in the level of financial resources available to a school district and the academic outcomes of the students. Importantly, we find this result using variation that is not derived from school financial reforms, and our results (which use data through 2015) relate to contemporaneity spending levels. Our results (based on spending cuts) are important because show that school districts respond to budget cut by disproportionately reducing non-core operational spending. However, our results do not support the notion that there is sufficient waste and ”bloat” in the public education system that budget cuts have no effect on outcomes. To the contrary, our results suggest that while school districts are able to offset cuts to core operations by delaying and abandoning construction projects, the scope for this is small in the face of large cuts. Unfortunately, we find students that experienced reduced public school spending had both lower test scores but also less high school completion. These patterns suggest that (a) school spending cuts do matter, and that (b) the ill-effects of the recession on the affected youth (through reduced public school spending) will be felt for years.

References

Elizabeth Oltmans Ananat, Anna Gassman-Pines, Dania Francis, and Christina Gibson-Davis. Children Left Behind: The Effects of Statewide Job Loss on Student Achievement. Technical report, National Bureau of Economic Research, Cambridge, MA, 6 2011.

D Mark Anderson and Mary Beth Walker. Does Shortening the School Week Impact Student Performance? Evidence from the Four-Day School Week. *Education Finance and Policy*, 10(3): 314–349, 2015. doi: 10.1162/EDFP{_}a{_}00165.

Matt Barnum. Christie Plan for Funding NJ Schools Widely Criticized; Cami Anderson Warns of 'Devastating Impact' — The 74, 2016.

Julian R. Betts. Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature. In Gary Burtless, editor, *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, pages 141–91. Brookings Institution, Washington D.C., 1996.

Christopher A. Candelaria and Kenneth A. Shores. Court-Ordered Finance Reforms in The Adequacy Era: Heterogeneous Causal Effects and Sensitivity. *Education Finance and Policy*, pages 1–91, 6 2017. ISSN 1557-3060. doi: 10.1162/EDFP{_}a{_}00236.

David Card and A. Abigail Payne. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, 83(1):49–82, 10 2002. ISSN 00472727. doi: 10.1016/S0047-2727(00)00177-8.

Rajashri Chakrabarti, Max Livingston, Elizabeth Setren, Jason Bram, Erica Groshen, Andrew Haughwout, James Orr, Joydeep Roy, Amy Ellen Schwartz, and Giorgio Topa. The Impact of the Great Recession on School District Finances: Evidence from New York. 2015.

James S. Coleman. Equality of Educational Opportunity, Reexamined. Technical report, Department of Education, 1990.

Noelle M. Ellerson. A Cliff Hanger : How America's Public Schools Continue to Feel the Impact of the Economic Downturn. Technical report, 4 2010.

William N. Evans, Robert M. Schwab, and Kathryn L. Wagner. The Great Recession and Public Education. *Education Finance and Policy*, pages 1–50, 9 2017. ISSN 1557-3060. doi: 10.1162/edfp{_}a{_}00245.

Maria D. Fitzpatrick and Michael F. Lovenheim. Early Retirement Incentives and Student Achievement. *American Economic Journal: Economic Policy*, 6(3):120–154, 8 2014. ISSN 1945-7731. doi: 10.1257/pol.6.3.120.

Philip Gigliotti and Philip Gigliotti. Education Expenditures and Student Performance: Evidence from the Save Harmless Provision in New York State. 11 2017.

Craig Gundersen, Brent Kreider, and John Pepper. The Economics of Food Insecurity in the United States. *Applied Economic Perspectives and Policy*, 33(3):281–303, 2011. doi: 10.1093/aepp/ppr022.

E. A. Hanushek. Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19(2):141–164, 1 1997. ISSN 0162-3737. doi: 10.3102/01623737019002141.

Paul T. Hill and Georgia Heyward. A troubling contagion: The rural 4-day school week, 2017.

Michael Hurd and Susann Rohwedder. Effects of the Financial Crisis and Great Recession on American Households. Technical report, National Bureau of Economic Research, Cambridge, MA, 9 2010.

Joshua Hyman. Does money matter in the long run? Effects of school spending on educational attainment. *American Economic Journal: Economic Policy*, 9(4):256–280, 11 2017. ISSN 1945774X. doi: 10.1257/pol.20150249.

C. Kirabo Jackson. What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, 5 2018. doi: 10.3386/w22226.

C. Kirabo Jackson, Jonah E. Rockoff, and Douglas O. Staiger. Teacher Effects and Teacher-Related Policies. *Annual Review of Economics*, 6(1):801–825, 8 2014. ISSN 1941-1383. doi: 10.1146/annurev-economics-080213-040845.

C. Kirabo Jackson, Rucker C. Johnson, and Claudia Persico. The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, 131(1):157–218, 2 2016. ISSN 0033-5533. doi: 10.1093/qje/qjv036.

Rucker C Johnson and C. Kirabo Jackson. Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending. *NBER Working Paper No. 23489*, No. 23489:1–61, 6 2017. doi: 10.3386/w23489.

Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263, 3 1979. ISSN 00129682. doi: 10.2307/1914185.

Greg Kaplan, Kurt Mitman, and Giovanni Violante. The Housing Boom and Bust: Model Meets Evidence. Technical report, National Bureau of Economic Research, Cambridge, MA, 8 2017.

Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics*, 0(0), 2018. ISSN 1945-7782. doi: 10.1257/APP.20160567.

Richard R. Lau. Negativity in political perception. *Political Behavior*, 4(4):353–377, 1982. ISSN 0190-9320. doi: 10.1007/BF00986969.

Edward P. Lazear, Kathryn L. Shaw, and Christopher Stanton. Making Do with Less: Working

Harder during Recessions. *Journal of Labor Economics*, 34(S1):S333–S360, 1 2016. ISSN 0734-306X. doi: 10.1086/682406.

Michael Leachman and Chris Mai. Most States Still Funding Schools Less Than Before the Recession. 2014.

Michael Leachman, Nick Albares, Kathleen Masterson, and Marlana Wallace. Most States Have Cut School Funding, and Some Continue Cutting. 2016.

Michael Leachman, Kathleen Masterson, and Eric Figueroa. A Punishing Decade for School Funding. 2017.

Jeffrey B. Liebman and Neale Mahoney. Do expiring budgets lead to wasteful year-end spending? Evidence from federal procurement, 11 2017. ISSN 00028282.

Corbin Miller. The Effect of Education Spending on Student Achievement: Evidence from Property Wealth and School Finance Rules. 2017.

Philip Oreopoulos, Marianne Page, and Ann Huff Stevens. The Intergenerational Effects of Worker Displacement. *Journal of Labor Economics*, 26(3):455–000, 7 2008. ISSN 0734-306X. doi: 10.1086/588493.

Diane Schanzenbach and Bauer Lauren. Food Insecurity among Children in Massachusetts, 2017. URL http://brook.gs/2s6Xzkz.

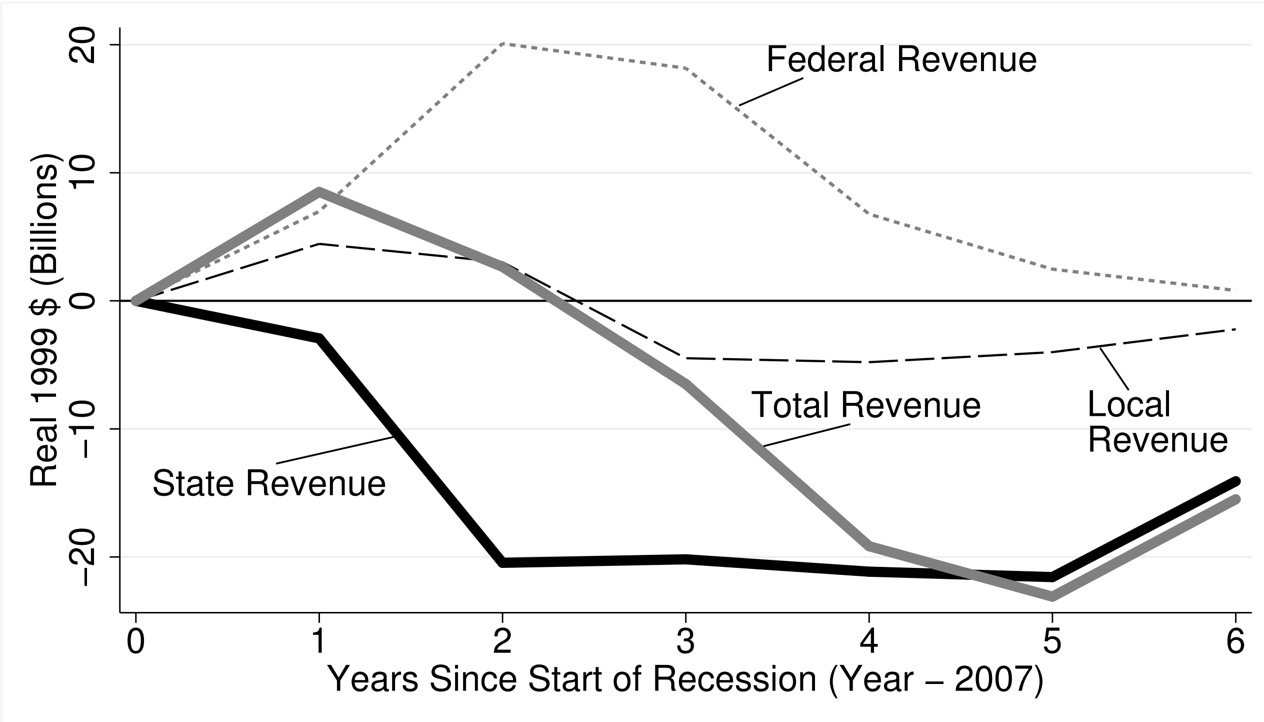
Kenneth Shores and Matthew Steinberg. The Impact of the Great Recession on Student Achievement: Evidence from Population Data. 8 2017.

Ann Huff Stevens and Jessamyn Schaller. Short-run effects of parental job loss on children's academic achievement. *Economics of Education Review*, 30(2):289–299, 2011.

Danny Yagan. Employment Hysteresis from the Great Recession. 9 2017.

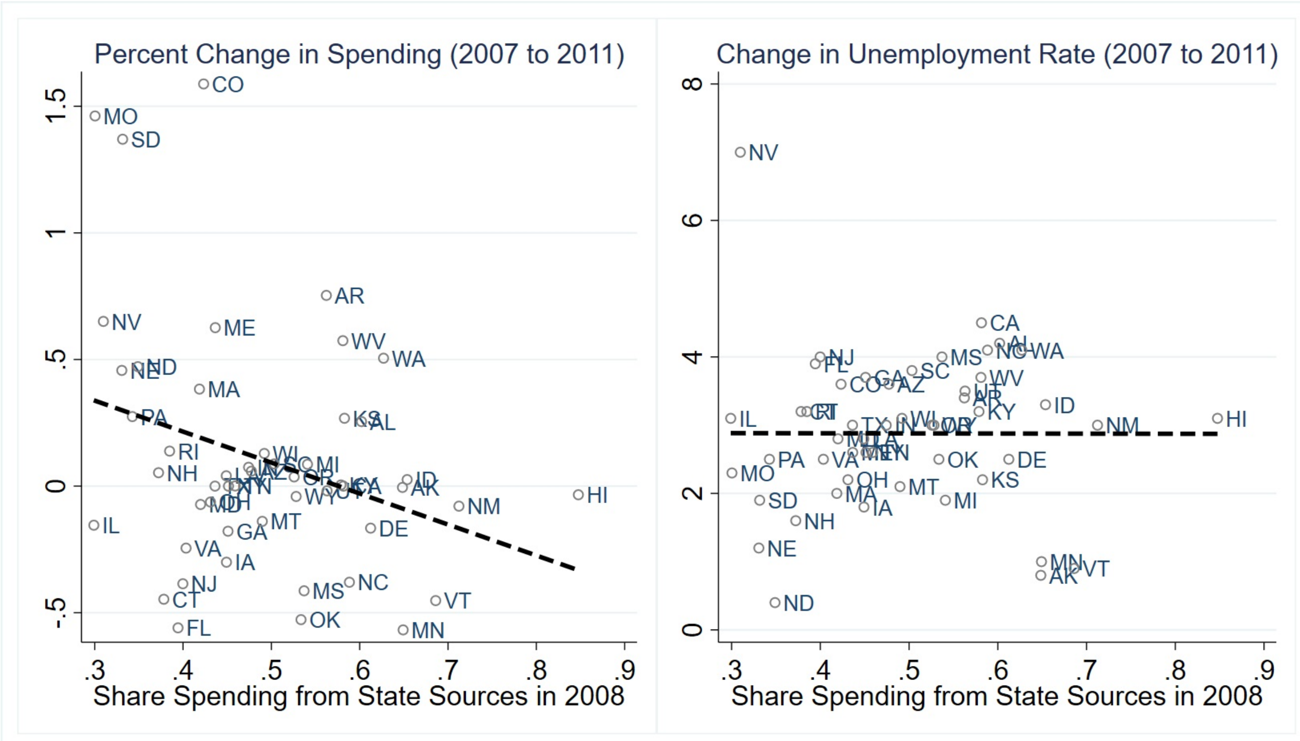
Tables and Figures

Figure 1. *Source of Revenues for K12 Education after the recession.*



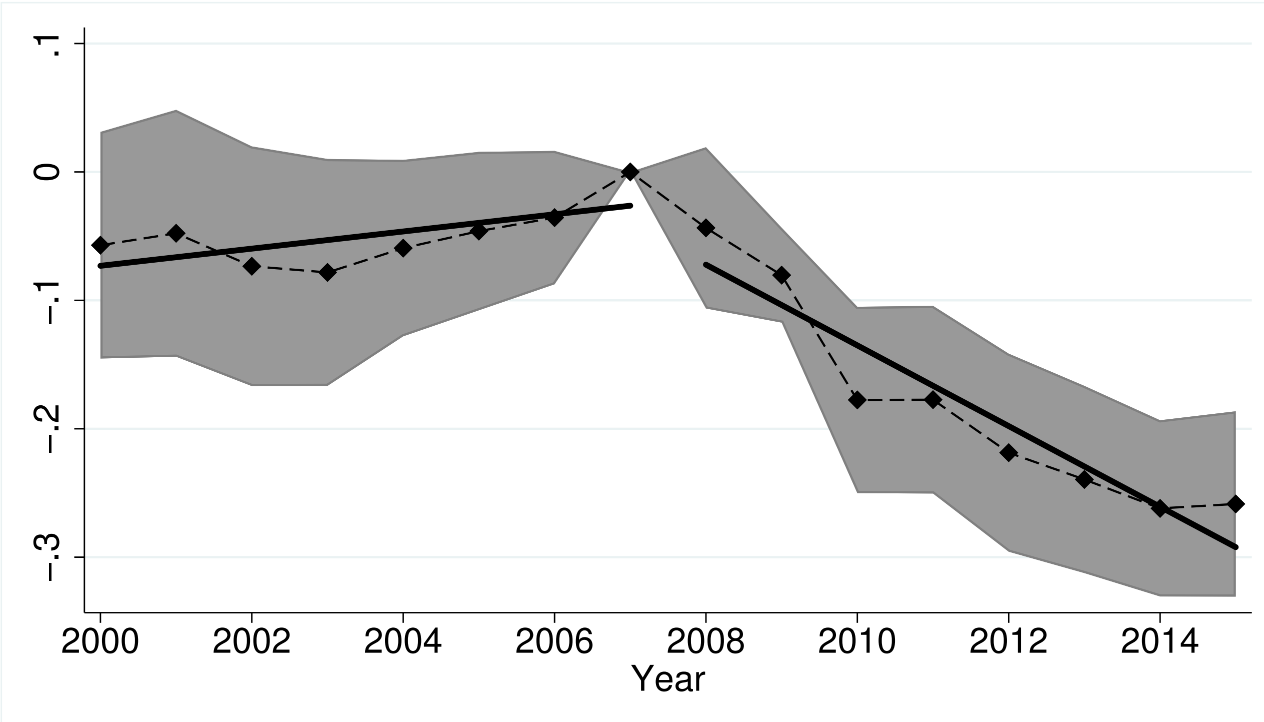
Notes: This figure plots the change in national aggregate revenue (summed over all available districts in the CCD data) for public schools relative to 2007 levels. The total revenue numbers are broken down by the source of funding (federal, state, and local); changes in each of which is also shown separately. Nominal dollars in all years are deflated by the CPI and adjusted to real 1999 dollars in billions. As a frame of reference, the values in 2007 are as follows: Total Revenue: 457.95 billions, State Revenue: 217.37 billions, Local Revenue: 204.15 billions, Federal Revenue: 36.42 billions.

Figure 2. *Fraction of K12 Revenue from State Sources: Spending Growth and Unemployment.*



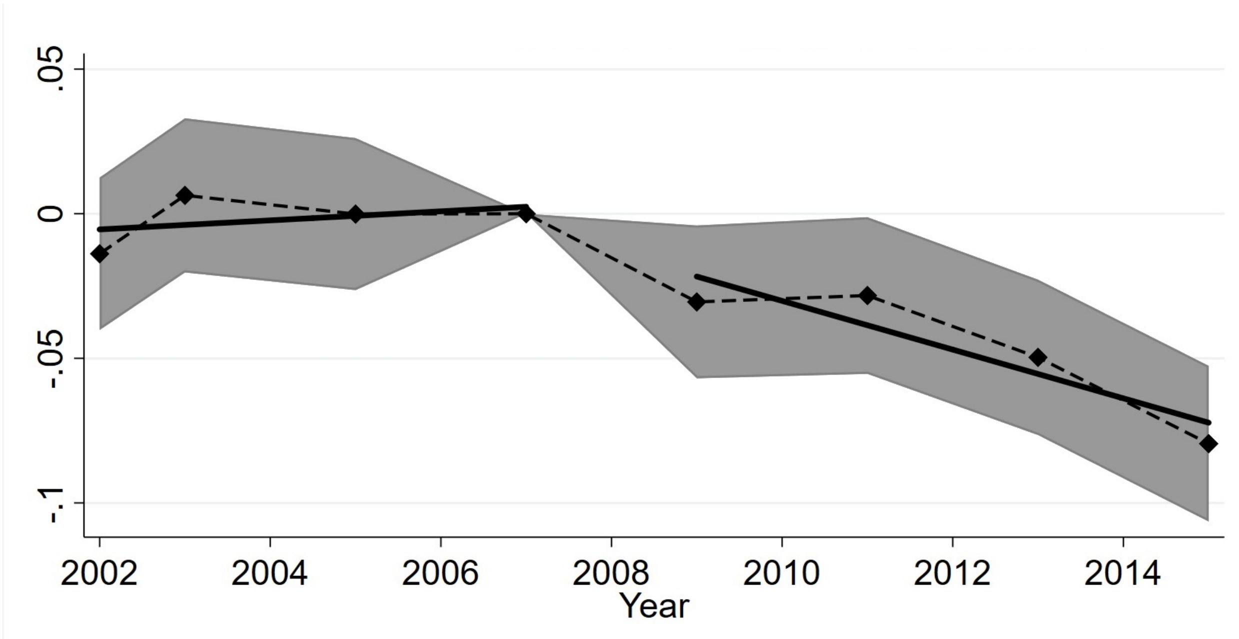
Notes: The left panel shows the percent change in state average (averaged across all districts in the state) K12 spending between 2007 and 2011 by the percent of state K12 revenues that came from state sources. The right panel shows the percent change in the state unemployment rate between 2007 and 2011 by the percent of state K12 revenues that came from state sources. These plots exclude D.C. because it is an outlier in the distribution of percent of state spending. Note that all regression results include all states.

Figure 3. *Difference in Log Spending Between States with High and Low Reliance on State Revenues Over Time*



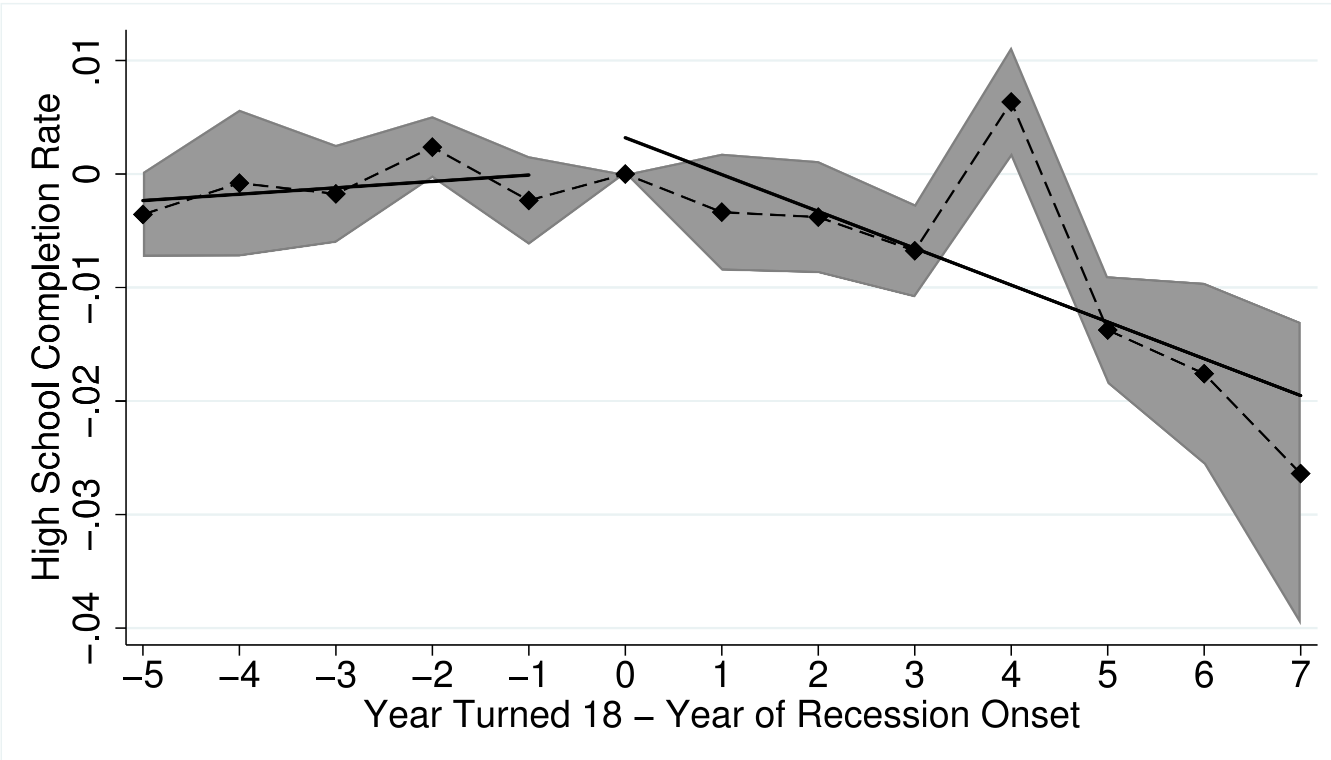
Notes: The dashed connected line depicts the coefficients on the individual calendar year indicators interacted with the state reliance on state revenue in 2008, $\ln\Omega_s$. More specifically, it plots the β_t for each calendar year from Equation 5) relative to the coefficient for 2007. The gray area depicts the 90 percent confidence interval for each individual calendar year interaction. The solid lines represent the linear fit during the pre-recession period (negative values of exposure) and post recession periods (non-negative values of exposure).

Figure 4. *Difference in NAEP Scores Between States with High and Low Reliance on State Revenues Over Time.*



Notes: The dashed connected line depicts the coefficients on the individual calendar year indicators interacted with the state reliance on state revenue in 2008, $\ln\Omega_s$. More specifically, it plots the β_t for each calendar year from Equation 6) relative to the coefficient for 2007. The gray area depicts the 90 percent confidence interval for each individual calendar year interaction. The solid lines represent the linear fit during the pre-recession period (negative values of exposure) and post recession periods (non-negative values of exposure).

Figure 5. *Difference in High-school Completion Rates Between Exposed Cohorts from States with High and Low Reliance on State Revenues: By Years of Recession Exposure.*



Notes: The dashed connected line depicts the coefficients on the individual recession exposure year indicators interacted with the state reliance on state revenue in 2008, $\ln\Omega_s$. More specifically, it plots the β_e for each exposure year from Equation 7) relative to the coefficient for cohorts who graduates in 2007. The gray area depicts the 90 percent confidence interval for each individual exposure year interaction. The solid lines represent the linear fit for the pre-recession cohorts (negative values of exposure) and post recession cohorts (non-negative values of exposure).

Table 1. Summary Statistics

	N	Mean	SD
Student Level (2002, 2003,2005,2007,2009,2011,2013,2015)			
NAEP Score (Standardized on 2003 scores)	4289000	0.0831	1.003
Limited English Proficiency	4289000	0.0661	0.248
School Lunch Eligibility	4289000	0.494	0.483
District Level (2002-2015)			
Total Population	193,190	24,099	108,049
Child-Aged Population	193,190	3,943	17,739
Child-Aged Population in Poverty	193,190	702.6	4,599
Public School Enrollment	193,190	3,488	14,765
Bartik Instrument (predicted unemplyment)	193,190	0.06327	0.01499
Annual Average Employment	193,190	119,232	296,863
District - Spending Per Student (2002-2015)			
Per Pupil Spending (in thousands)	193,190	10.07	5.594
Elementary/Secondary Spending	192,650	8,340	3,551
Instructional Spending	192,809	5,035	2,111
Support Services Spending	192,550	2,953	1,530
Other Spending	192,876	343.7	180.6
Capital Expenditures	192,771	821.5	1,480
Construction Expenditures	192,783	572.8	1,358
Non-Construction Expenditures	193,193	259.7	1,182
Non Elementary/Secondary Spending + Payments	193,193	815.7	3,179
Total Salaries	192,669	4,859	1,908
Instructional Salaries	192,863	3,330	1,295
Non-Instructional Salaries	193,193	1,601	2,416
Total Benefits	192,830	1,456	896.8
District - Staffing Per Student (2002-2015)			
Teachers	187,744	0.076	0.0271
Aides	168,138	0.0195	0.0149
Guidance Counselors	155,664	0.00261	0.00139
Library Staff (Specialists + Support Staff)	126,282	0.00295	0.00198
LEA Staff (Administrators + Support Staff)	156,602	0.00765	0.00722
State Level (2002 -2015)			
IV (%State Share in 2008)	714	0.493	0.135
Unemployment Rate	714	0.0615	0.0202
Bartik Instrument	714	0.0616	0.0148
State by Birth Cohort by Year Level (1994 - 2015)			
Age	4500	22.41	3.272
High School Completion Rate	4500	0.89	0.0976

Notes: See Appendix VI for details on each data source.

Table 2. First Stage Regressions. Districts with NAEP Scores

	1	2	3	4
	First Stage Regressions			
	Per Pupil Spending/1000		Log of Per Pupil Spending	
Year*ln(%State$_{2008}$)	0.0167	0.0176	0.00469**	0.00478**
	[0.0149]	[0.00946]	[0.00126]	[0.000412]
Post*ln(%State$_{2008}$)	491.5**	1,938**	42.35**	167.3**
	[106.6]	[60.80]	[8.322]	[4.996]
Post*Year*ln(%State$_{2008}$)	-0.245**	-0.967**	-0.0211**	-0.0835**
	[0.0530]	[0.0306]	[0.00414]	[0.00253]
(Not HI)*Post*ln(%State$_{2008}$)		-1,432**		-123.8**
		[67.16]		[5.523]
(Not HI)*Post*Year*ln(%State$_{2008}$)		0.715**		0.0618**
		[0.0331]		[0.00272]
Observations	4,260,116			
First Stage F Statistic	34.55	270245	25.29	128210
Prob>F	0.000236	<0.0001	0.000627	<0.0001

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year.

** p<0.01, * p<0.05, + p<0.1

All models include year and district fixed effects.

Sample is all districts with NAEP data. Data are at the individual student level.

Table 3. Exogeneity of Spending: Districts with NAEP Scores

	1	2	3	4	5
	2SLS Regressions				
	State Unemployment	County Child Poverty	Log County Child Poverty	County Employment	Log County Employment
	Panel 1: Controling for total district population and population of school-age children with district and year fixed effects, state and district bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007.				
Log Per Pupil Spending	0.00528	524.7	-1.698+	61,221	0.154
	[0.00779]	[5,704]	[0.860]	[48,306]	[0.120]
	Panel 2: All controls in Panel 1 with additional controls for linear time trend in $\ln(\%State_{2008})$				
Log Per Pupil Spending	0.00374	-750.9	-2.249	-31,927	-0.0305
	[0.0153]	[6,550]	[1.552]	[66,741]	[0.132]
$Year*\ln(\%State_{2008})$	1.32E-06	-10.57	-0.00818	-1,272+	-0.00253+
	[0.000192]	[63.58]	[0.0125]	[580.1]	[0.00126]
Observations			4,260,116		

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** p<0.01, * p<0.05, + p<0.1

All models include district and year fixed effects and control for grade, subject, LEP, School Lunch eligibility, and district total and child-aged population.

Sample is all districts with NAEP data. Data are at the individual student level.

Table 4. Two-Stage-Least-Squares Regressions: School Spending Categories

	1	2	3	4	5	6	7	8	9	10	11
	Expenditure Categories (Per Pupil)										
	Total Expenditures			Capital Spending		Elementary/Secondary Spending					
	Capital	Elem/Sec	Other [a.]	Construction	Non-Construction	Instructional	Support Services	Other	Instructional Salaries	Non-Instructional Salaries	Benefits
Mean (Dependent Var.)	821.5	8340	815.7	572.8	259.7	5035	2953	343.7	3330	1601	1456
Average Share of Spending	0.0738	0.864	0.062	0.0491	0.0247	0.523	0.303	0.0385	0.35	0.16	0.149
Per Pupil Spending	0.449**	0.554**	0.0608	0.466**	-0.0165	0.439**	0.0768	0.0365	0.201**	-0.0938	0.103+
	[0.138]	[0.131]	[0.0471]	[0.129]	[0.0331]	[0.101]	[0.0982]	[0.0244]	[0.062]	[0.0657]	[0.0521]
Year*(%State_2008)	54.90**	-48.19*	-6.534	61.23**	-6.359	-41.99*	-11.03	4.599	-18.83*	-17.28+	-8.428
	[16.21]	[18.74]	[9.562]	[16.49]	[4.653]	[19.12]	[12.85]	[3.612]	[7.899]	[8.212]	[17.00]
P(Average=Marginal)	0.0173	0.035	0.256	0.00678	0.235	0.419	0.0388	0.937	0.0319	0.00195	0.398
Observations	193190										

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** $p<0.01$, * $p<0.05$, + $p<0.1$

All models control for total district population and population of school-age children with district and year fixed effects, state and district Bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007 and additional controls for linear time trend in $\ln(\%State_{2008})$.

Sample is all available districts. Regression models are weighted by Public School Enrollment.

a. Other spending includes Non Elem/Sec and Payments.

Table 5. Two-Stage-Least-Squares Regressions: Staffing Categories

	1	2	3	4	5
	Staffing Categories				
	Teachers Per Pupil	Aides Per Pupil	Guidance Counselors Per Pupil	Library Staff Per Pupil	LEA Staff Per Pupil
Log Per Pupil Spending	0.0437*	-0.0358	0.00165	-0.00146	-0.00771
	[0.0182]	[0.0345]	[0.00155]	[0.00246]	[0.0158]
Year*ln(%$State_{2008}$)	-6.33E-05	-0.000781	9.65E-06	-3.66E-05	-0.000262
	[0.000308]	[0.000567]	[2.48e-05]	[2.30e-05]	[0.000216]
Mean(Dependent Var)	0.076	0.0195	0.00261	0.00295	0.00765
Observations	187744	177964	176884	137061	173701

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** $p<0.01$, * $p<0.05$, + $p<0.1$

All models control for total district population and population of school-age children with district and year fixed effects, state and district bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007 and additional controls for linear time trend in ln(%$State_{2008}$). Sample is all available districts. Regression models are weighted by Public School Enrollment.

Table 6. Two-Stage-Least-Squares Regressions. Districts with NAEP Scores

	1	2	3	4	5	6	7	8
	NAEP Score (2002-2015)							
Per Pupil Spending/1000	0.0604**	0.0648**	0.0705**	0.0710*				
	[0.0160]	[0.0137]	[0.0165]	[0.0241]				
Log Per Pupil Spending					0.910**	0.968**	0.782**	0.783**
					[0.207]	[0.175]	[0.110]	[0.213]
Year*ln(%State$_{2008}$)			0.000935	0.00146			-0.00228	-0.00179
			[0.00290]	[0.00364]			[0.00229]	[0.00260]
Basic Contols	Y	Y	Y	Y	Y	Y	Y	Y
Full Controls		Y	Y	Y		Y	Y	Y
Full Controls and Economic Variables				Y				Y
Observations	4,260,116							

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** p<0.01, * p<0.05, + p<0.1

All model control for total district population and population of school-age children with district and year fixed effects (Basic controls), student grade, subject, LEP and Free/Reduced Lunch eligibility. Full controls also include state and district Bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007. Economic variables include the state unemployment rate, the district average annual employment and the number of children living in poverty in the district.

Sample is all available NAEP observations linked to districts.

Table 7. Two-Stage-Least-Squares Regressions: By Subject and Grade.

	1	2	3	4	5	6
Grade	All	All	Fourth	Fourth	Eighth	Eighth
Subject	Math	Reading	Math	Reading	Math	Reading
Log Per Pupil Spending	1.183**	0.355**	1.009**	0.292**	1.455*	0.465**
	[0.317]	[0.0503]	[0.184]	[0.0427]	[0.472]	[0.115]
Observations	2,012,813	2,247,109	1,076,596	1,205,595	936,152	1,041,438

All models control for total district population and population of school-age children with district and year fixed effects student grade, subject, LEP and Free/Reduced Lunch eligibility, state and district Bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007 and additional controls for linear time trend in $\ln(\%State_{2008})$.

Sample is all available NAEP observations linked to districts.

Table 8. 2SLS Regressions: Public and Private School State Averages

	1	2	3	4	5	6	7	8
	Mean NAEP Score							
Log Per Pupil Spending	0.728*	0.667*	0.725+	0.726+				
	[0.308]	[0.297]	[0.409]	[0.411]				
Log Public Per Pupil Spending					0.796*	0.892**	0.818*	0.839*
					[0.309]	[0.302]	[0.407]	[0.411]
Log Per Pupil Spending* Private					0.187	0.254	0.203	0.24
					[0.405]	[0.408]	[0.487]	[0.492]
Year*ln(%$State_{2008}$)			-0.00033	-0.00072			-0.00012	-0.00031
			[0.00384]	[0.00368]			[0.00383]	[0.00368]
Full Controls	N	Y	Y	Y	N	Y	Y	Y
Economic Controls	N	N	N	Y	N	N	N	Y
Pr(private=public)	-	-	-	-	0.0413	0.0325	0.0376	0.0419
Observations	730							

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** p<0.01, * p<0.05, + p<0.1

All model control for total district population and population of school-age children with district and year fixed effects (Basic controls) and average student grade, subject, LEP and Free/Reduced Lunch eligibility. Full controls also include the Bartik instrument and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007 and additional controls for linear time trend in ln(% $State_{2008}$). Economic variables include the state unemployment rate, the district average annual employment and the number of children living in poverty in the district.

Table 9. Two-Stage-Least-Squares Regressions: High School Completion Rates (IPUMS)

	1	2	3	4	5	6	7	8
	High School Completion Percentage (0/1)							
Log (Per pupil spending:ages 15 to 18)	0.242*	0.288*	0.277*	0.261**				
	[0.0910]	[0.116]	[0.107]	[0.0947]				
Log (Per pupil spending: ages 6 to 18)					0.474**	0.412*	0.716+	0.739+
					[0.167]	[0.161]	[0.364]	[0.377]
Cohort*ln(%State$_{2008}$)			-0.000124	-0.000283			0.00174	0.00146
			[0.000295]	[0.000352]			[0.00130]	[0.00132]
Full Controls	N	Y	Y	Y	N	Y	Y	Y
Economic Controls	N	N	N	Y	N	N	N	Y
Observations	4,500				4,005			
R-squared	0.779	0.78	0.781	0.783	0.789	0.795	0.779	0.778

Notes: Robust standard errors in brackets adjusted for two-way clustering by state and year

** $p<0.01$, * $p<0.05$, + $p<0.1$

All models control for total district population and population of school-age children with district and year fixed effects (Basic controls). Full controls also include state and district Bartik instruments and time fixed effects interacted with the state unemployment in 2007 and district child poverty rate in 2007. Economic variables include the state unemployment rate, the district average annual employment and the number of children living in poverty in the district.

Appendix

A Data on School Spending and Resources

The data on district level school finances is collected from the Census website.[17] The underlying data comes from the Common Core of Data (CCD) School District Finance Survey (F-33). It consists of data submitted annually to the National Center for Education Statistics (NCES) by state education agencies (SEAs) in the 50 states and the District of Columbia. The purpose of the survey is to provide finance data for all local education agencies (LEAs) that provide free public elementary and secondary education in the United States. Both NCES and the Governments Division of the U.S. Census Bureau collect public school system finance data, and they collaborate in their efforts to gather these data. The F-33 data provides information on revenues, expenditures, and the number of students enrolled. Expenditures are reported in a number of categories including instructional spending, capital outlays, and administrative spending. Revenues are reported in several fine categories and aggregated to local, state, and federal sources.

The surveys are administered annually from 1992 onward. We link together multiple years of data to create a balanced panel school district data set for analysis. In constructing the data set, we found that the financial data contained some extremely large and small values. These values could be valid, but it is more likely that some districts incorrectly reported enrollments or expenditures. We therefore censored the data by deleting extreme values. First, we calculated the (unweighted) 99th and 1st percentile district in total per-pupil current expenditures for each state and year. We then capped values of districts with per-pupil expenditures at greater than 200 per-cent of the 99th percentile of per-pupil revenues or less than 50 percent of the 1st percentile.

For school spending categories (such as capital or instructional salaries), we replace values with missing where the cpi-adjusted per-student categorical spending value is more than twice the 99th percentile. We follow a similar strategy for reported staffing categories, which come from the NCES Common Core of Data LEA Universe surveys (CCD), replacing staffing values with missing if the total staffing variable or the staffing per student (or students per staff) is more than twice the 99th percentile.

B Recession Intensity & Employment Data

Important to our identification strategy is controlling for the direct effect of broader recessionary economic conditions. For this purpose, we construct an index of recession severity and exposure. We exploit the fact that the impact of recession varied on basis of local industrial compositions and create a shift-share instrument, along the lines of Bartik (1991), which captures changes in economic conditions attributable to the onset of the recession.

To do so, we follow the steps broadly outlined in Yagan (2017). We retrieve average annual county-level employment data from the Quarterly Census of Employment and Wages (QCEW).[18] Each county's time-varying shift-share shock is computed as the projected unemployment in each year, based on the interaction between the 2007, pre-recession, employment composition by two-digit NAICS industry categories and the nationwide unemployment by the same groupings in that year. Formally, in county c during year t the instrument equals:

[17] For instance, data for the school year 2015-2016 is available at https://www.census.gov/data/tables/2015/econ/school-finances/secondary-education-finance.html and data for the other years can be retrieved by modifying the appropriate part of the url.

[18] The QCEW program publishes an annual count of employment and wages reported by employers covering 98 percent of U.S. jobs, available at the county, MSA, state and national levels by industry. Average annual data were downloaded from the Bureau of Labor Statistics for each county and year from https://www.bls.gov/cew/datatoc.htm

$$Bartik\ Predictor_{ct} = \sum_j \left(\frac{E_{jc2007}}{\sum_{j'} E_{j'c2007}} \times \text{National Unemployment}_{jct} \right) \tag{10}$$

where j denotes a two-digit industry, E_{jc2007} denotes total employment in industry j in county c in 2007, and $National\ Unemployment_{jt}$ is the nationwide unemployment rate in industry j in county c in year t.

From the same dataset (QCEW) above, we also compile the annual total employment number in each county as an additional measure of economic status.

To utilize these controls in the analysis, we merged the county level data on recession shock to schools to the data base to district level data from the Common Core and F33 outlined above. It is not straight forward to match the records since the district to county mapping is not one to one. To overcome this complication, we use a spatial correspondence file from http://mcdc.missouri.edu/websas/geocorr12.html. This file identifies the exact percentage of a district's area contained in each county. Using this information, we link the district level observations to the weighted average of values in the counties that the district overlaps. The resulting value approximates the shift-share predictor in the district, assuming industrial compositions are distributed uniformly within each county.

C NAEP Data

We use restricted-use individual-level NAEP test score data from the National Center for Education Statistics. We infix the raw files to Stata, including all plausible score values per student, and restricting the sample to the NAEP reporting sample and public school students, except in the case of our robustness test, which includes private school students as well. The restriction to the reporting sample and public school students corresponds directly to the sample used to calculate state averages as reported publicly by NCES.

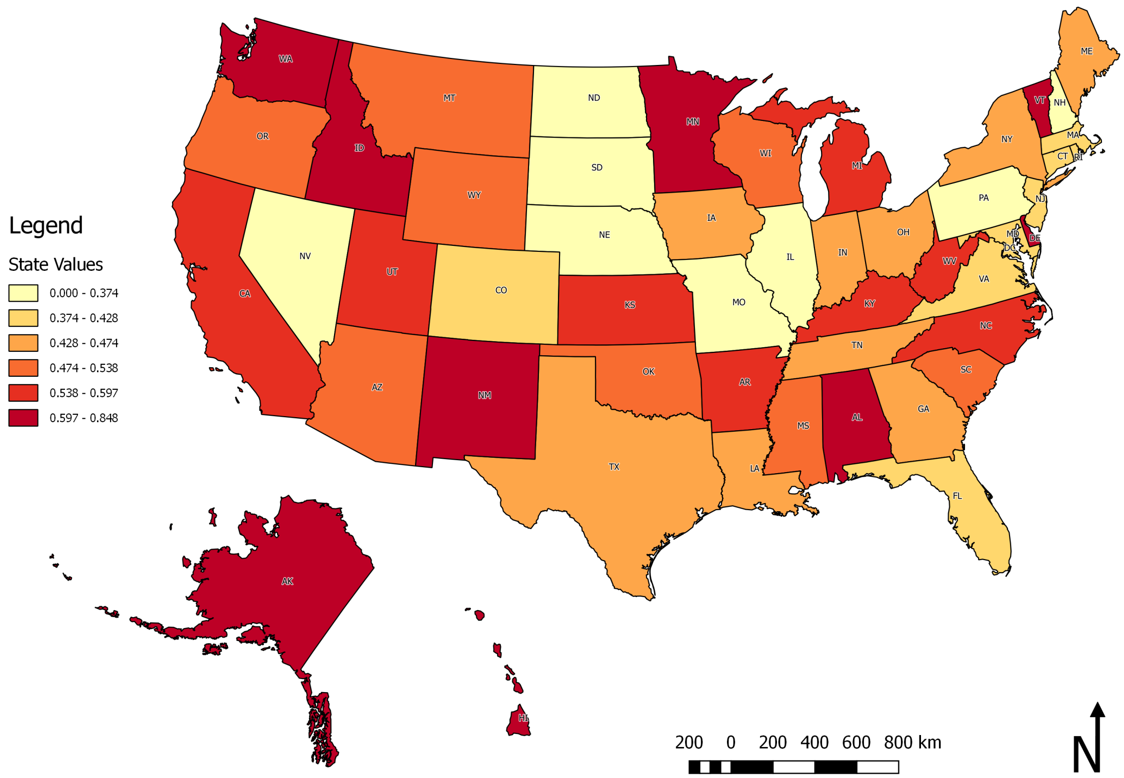
We link student scores to school districts using the reported NCES school identifier. Our dependent variable in all NAEP estimations is the mean of all reported plausible values for each student, year, grade, and subject, standardized to the base year of 2003. We restrict our analyses to the years 2002 and later, as NAEP sampling increased dramatically after 2001 and testing years became more consistent at this time. In data that are aggregated to the state level, we weight the state-averages by the reported original weight, ORIGWT.

Appendix Table 1: NAEP Availability

Year	4th Grade Math	8th Grade Math	4th Grade Reading	8th Grade Reading	Tested Students
2015	X	X	X	X	430438
2014					
2013	X	X	X	X	575298
2012					
2011	X	X	X	X	619789
2010					
2009	X	X	X	X	571308
2008					
2007	X	X	X	X	620220
2006					
2005	X	X	X	X	589458
2004					
2003	X	X	X	X	642244
2002			X	X	240228
2001					
2000	X	X	X		22246
1999					
1998			X	X	15391
1997					
1996	X	X			10805
1995					
1994			X		6030
1993					
1992	X	X	X		16719

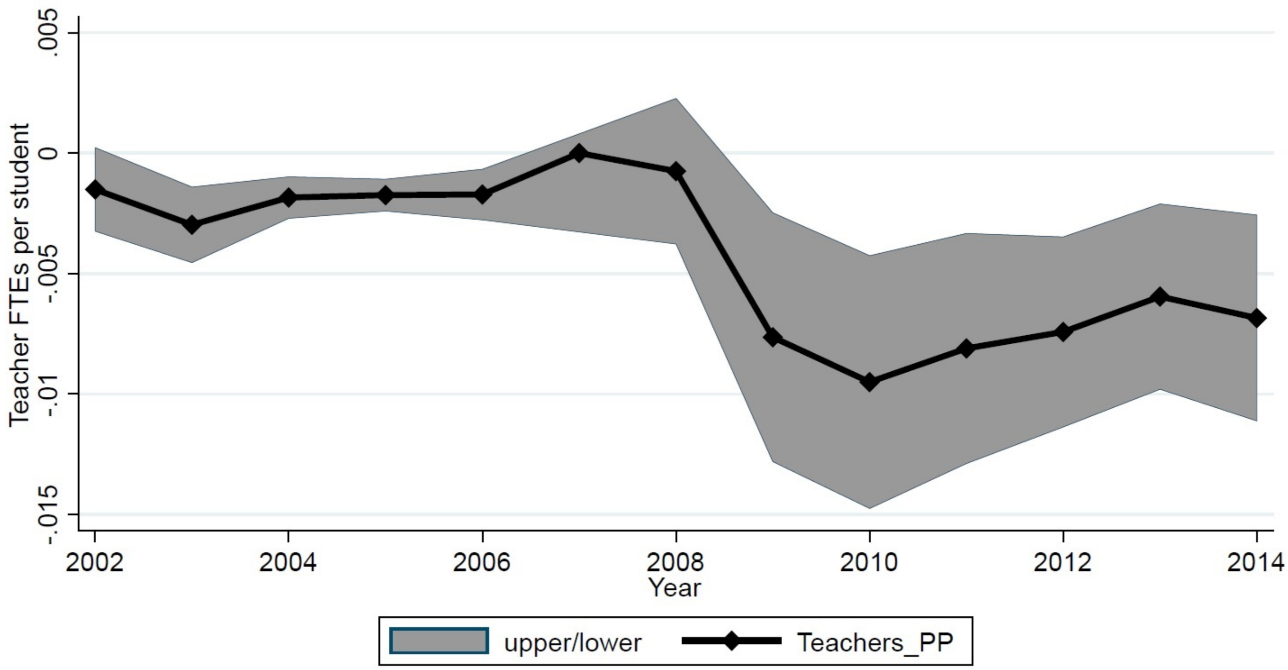
Notes: NAEP availability by subject-grade in sample.

Appendix Figure 1. Fraction of Education Revenue from State Sources



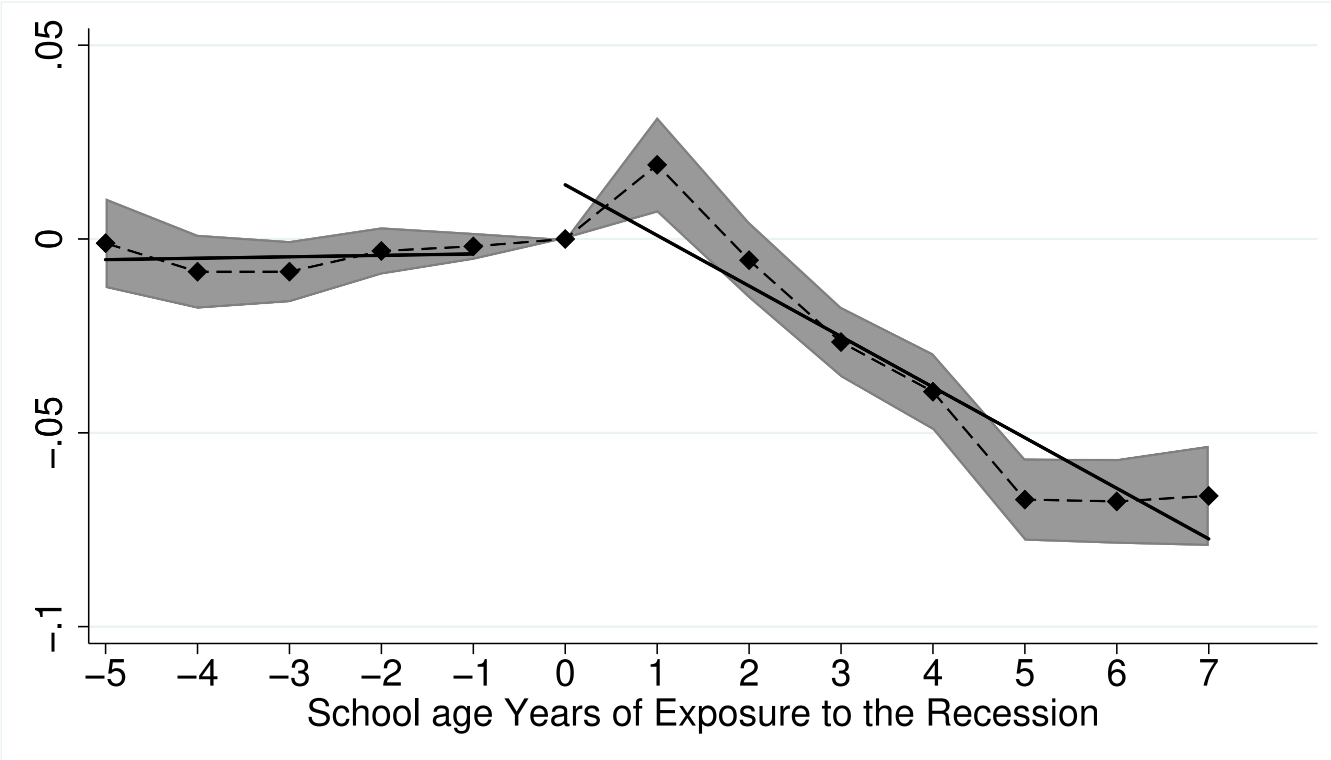
Notes: This map shows the extent of variation in $(\%)State_s$ across the country.

Appendix Figure 2. Marginal Effect of IV on Staffing Ratio.
All Districts



Notes: The dashed connected line depicts the coefficients on the individual calendar year indicators interacted with the state reliance on state revenue in 2008, $\ln\Omega_s$. More specifically, it plots the β_t for each calendar year from Equation 6) relative to the coefficient for 2007. The gray area depicts the 90 percent confidence interval for each individual calendar year interaction. The solid lines represent the linear fit during the pre-recession period (negative values of exposure) and post recession periods (non-negative values of exposure).

Appendix Figure 3. Marginal Effect of IV by Year on Per-Pupil Spending (Between ages 15 to 18)



Notes: The connected dashed lines represents the individual exposure year interaction coefficients (relative to the effect for 2007). The gray area depicts the 90 percent confidence interval for each individual exposure year interaction. The solid lines represent the linear fit during the pre-recession period (negative values of exposure) and post recession periods (non-negative values of exposure).