



Beyond the School Building: Examining the Association Between Out-of-School Factors and Multidimensional School Grades

Nandrea Burrell



Erica Harbatkin

Florida State University
United States

Citation: Burrell, N., & Harbatkin, E. (2024). Beyond the school building: Examining the association between out-of-school factors and multidimensional school grades. *Education Policy Analysis Archives*, 32(32). <https://doi.org/10.14507/epaa.32.8497>

Abstract: Many states report school performance grades as a way to inform the public about school quality. However, past research has shown that when these grades drew largely on proficiency-based measures, they served to capture variation in school and community demographics rather than school quality. We extend this literature by examining whether a multidimensional measure of school quality such as those required under the Every Student Succeeds Act is less confounded by out-of-school factors than the proficiency measures that characterized previous generations of accountability. Drawing on school accountability grades from Florida combined with school and community demographic data, we find that more than half the variation in multidimensional measures of school quality can be explained by observable school- and county-level factors outside the school's locus of control. Together, our findings show that even school grades that draw on multiple measures misattribute the contribution of demographics and socioeconomics to school quality—but subcomponents based on learning gains perform better than those based on proficiency. We conclude with policy implications and recommend that states focus public reporting

on school quality measures driven less by out-of-school factors and more by the school's true contribution to student outcomes.

Keywords: accountability; educational indicators; education policy

Más allá del edificio escolar: Examen de la asociación entre factores extraescolares y resultados escolares multidimensionales

Resumen: Muchos estados informan las calificaciones de desempeño escolar como una forma de informar al público sobre la calidad de la escuela. Sin embargo, investigaciones anteriores han demostrado que cuando estos grados se basaron en gran medida en medidas basadas en el dominio, sirvieron para capturar la variación en la demografía de la escuela y la comunidad en lugar de la calidad de la escuela. Ampliamos esta literatura examinando si una medida multidimensional de la calidad de la escuela, como las requeridas por la Ley Every Student Succeeds, está menos confundida por factores extraescolares que las medidas de competencia que caracterizaron a las generaciones anteriores de rendición de cuentas. Basándonos en las calificaciones de responsabilidad escolar de Florida combinadas con datos demográficos de escuelas y comunidades, encontramos que más de la mitad de la variación en las medidas multidimensionales de calidad escolar puede explicarse por factores observables a nivel de escuela y condado fuera del locus de control de la escuela. En conjunto, nuestros hallazgos muestran que incluso los grados escolares que se basan en múltiples medidas atribuyen erróneamente la contribución de la demografía y la socioeconomía a la calidad de la escuela, pero los subcomponentes basados en avances en el aprendizaje obtienen mejores resultados que aquellos basados en el dominio. Concluimos con las implicaciones políticas y recomendamos que los estados centren los informes públicos en medidas de calidad escolar que se basen menos en factores extraescolares y más en la verdadera contribución de la escuela a los resultados de los estudiantes.

Palabras-clave: rendición de cuentas; indicadores educativos; política educativa

Além do edifício escolar: Examinando a associação entre fatores fora da escola e resultados escolares multidimensionais

Resumo: Muitos estados relatam notas de desempenho escolar como forma de informar o público sobre a qualidade da escola. No entanto, pesquisas anteriores demonstraram que, quando estas notas se baseavam em grande parte em medidas baseadas na proficiência, serviam para captar a variação na demografia escolar e comunitária, e não na qualidade da escola. Ampliamos esta literatura examinando se uma medida multidimensional da qualidade escolar, como as exigidas pela Lei de Every Student Succeeds, é menos confundida por fatores externos à escola do que as medidas de proficiência que caracterizaram as gerações anteriores de responsabilização. Com base nas notas de responsabilização escolar da Florida, combinadas com dados demográficos escolares e comunitários, descobrimos que mais de metade da variação nas medidas multidimensionais da qualidade escolar pode ser explicada por fatores observáveis a nível da escola e do condado, fora do locus de controlo da escola. Em conjunto, as nossas conclusões mostram que mesmo as notas escolares que se baseiam em múltiplas medidas atribuem erroneamente a contribuição da demografia e da socioeconomia para a qualidade da escola – mas os subcomponentes baseados nos ganhos de aprendizagem têm um desempenho melhor do que aqueles baseados na proficiência. Concluimos com implicações políticas e recomendamos que os estados concentrem os relatórios públicos em medidas de qualidade escolar que sejam impulsionadas menos por fatores externos à escola e mais pela verdadeira contribuição da escola para os resultados dos alunos.

Palavras-chave: *accountability*; indicadores educacionais; política educacional

Beyond the School Building: Examining the Association Between Out-of-School Factors and Multidimensional School Grades

State systems for assigning letter grades to schools have long been criticized for penalizing schools for the socioeconomic status of the student body rather than their effectiveness at supporting and educating the students they serve (Darling-Hammond, 2007; Figlio & Loeb, 2011; Lee & Reeves, 2012). These criticisms ramped up during No Child Left Behind (NCLB), which sought to improve educational opportunities for all students in part through a focus on required reporting of subgroup achievement. However, shining a light on student achievement shortfalls came with new concerns; in particular, that labeling schools as failing could lead to families moving or transferring away from schools, loss of local control and further disenfranchisement of already marginalized communities, and new challenges recruiting and retaining teachers due to the stigma of the failing grade (Darling-Hammond, 2007; Fusarelli, 2004; Gamoran, 2008, 2015; Harbatkin et al., 2024; Houston & Henig, 2023; Kim & Sunderman, 2005; Owens & Sunderman, 2006; Reardon, 2019). More than 20 years later, there is evidence that some but not all of these fears were borne out, as predominantly economically disadvantaged and Black communities, respectively, were disproportionately subject to accountability-driven takeovers, educators reported demoralization arising from the failing label, and economically disadvantaged students in at least one state may have been pushed out of the school system or reclassified by districts seeking to improve their scores (Gregg & Lavertu, 2023; Kitzmiller, 2020; Lipman, 2017; Pearman & Marie Greene, 2022; Strunk et al., 2016). On the other hand, there is evidence that schools receiving low accountability marks received needed resources for improvement (Dee et al., 2013), and in many cases they experienced sometimes sizeable achievement gains (Bonilla & Dee, 2020; Carlson & Lavertu, 2018; Dee & Jacob, 2011; Sun et al., 2017; Sun et al., 2021). Students subject to low pre-NCLB school accountability grades even experienced longer-term benefits in the form of higher educational attainment and lower adult criminal involvement and reliance on social welfare programs (Eren et al., 2023; Mansfield & Slichter, 2021).

Thus, the tradeoffs associated with school accountability policy are significant. Inherent in those tradeoffs is a question of the purpose of accountability policy—is it to name and shame educators and educational leaders into making improvements, or to provide additional support for schools most in need of additional resources (Darling-Hammond, 2007; Darling-Hammond & Snyder, 2015; Harbatkin & Wolf, 2023; Ladd, 2017)? There are arguments in favor of both theories of action. The theory of accountability underscores that transparency and measurement of school outcomes applies pressure on policy makers and educators to make necessary changes to improve, and also allows families to sort into better-performing neighborhoods and schools (Figlio & Loeb, 2011; Finnigan & Gross, 2007). School ratings on their own—especially different dimensions of school ratings—can also serve to provide information that is relevant to families with their own sets of educational priorities and goals (Burgess & Greaves, 2013). Thus, to the extent that school grades measure school quality, they can improve school systems and ultimately student achievement, both at the system level and for individual students whose families leverage the information in those grades to select into a school that addresses their unique needs. Indeed, there is evidence that grades on their own have induced meaningful change even without other interventions (Reback, 2008; Rouse et al., 2013; Winters & Cowen, 2012). There is also strong and growing evidence that increasing resources for underresourced schools can improve student outcomes (Candelaria & Shores, 2019; Jackson et al., 2016; Jackson & Mackevicius, 2024).

Thus, providing resources serves to benefit schools that need them, but penalizing schools with failing grades due to factors outside their control can undercut the goal of school accountability

by damaging the very schools it seeks to help. Under the Every Student Succeeds Act (ESSA, 2015), states are required to assign annual scores to schools using a multidimensional index of school quality that includes student achievement, growth, a non-academic measure or measures of school quality and student success, and other factors. Many states also use this index, or part of it, to assign letter grades as part of a state accountability system (Education Commission of the States, 2021). However, most of what we know thus far about letter grades is from the NCLB era and before, when grades were based largely on proficiency rates rather than student growth and other factors within the school system's locus of control. This research shows that school grades disproportionately penalize schools serving large shares of economically disadvantaged students and students from underrepresented minority groups, respectively, and that school grades do not adequately differentiate school quality (Adams, Forsyth, Ware, & Mwavita, 2016; Adams, Forysth, Ware, Mwavita, Barnes et al., 2016).

There is reason to believe that school grading systems based on ESSA's multidimensional school quality index could better differentiate schools based on quality rather than student demographic and socioeconomic background (Harbatkin & Wolf, 2023). We aim to provide evidence to examine this question using school accountability data from Florida, one of the first states in the U.S. to implement a consequential school accountability system (Figlio & Loeb, 2011). Specifically, we ask:

1. To what extent do different components of Florida's school accountability system appear to predict school quality over and above school and neighborhood race and socioeconomic factors?
2. To what extent does Florida's multidimensional school grade appear to predict school quality over and above school and neighborhood race and socioeconomic factors?

We answer these questions drawing on data from a snapshot in time after students throughout the nation returned to a new normal following the COVID-19 pandemic. Understanding the role of out-of-school factors in school grades is critical against the backdrop of the pandemic because research shows that schools undergoing accountability-driven reforms due to low performance and the communities they serve experienced some of the pandemic's most damaging effects (Cyrus et al., 2020; Finch & Hernández Finch, 2020; Harbatkin, Strunk, et al., 2023). We draw on publicly available 2022 school accountability data from the Florida Department of Education (FLDOE), the National Center for Education Statistics (NCES) Common Core of Data (CCD), and 2018-2022 five-year county-level estimates from the U.S. Census American Community Survey (ACS). We then predict school grades as a function of out-of-school factors related to student and community demographics and poverty to examine the extent to which school grades and their various components are explained by these observable factors. After capturing the contributions of these factors, some degree of the remaining variation could plausibly be explained by school quality—though it is certainly the case that there are unobserved out-of-school factors not included in our models and we are likely to be understating their contribution to school grades.

While there is no singular, agreed-upon definition of school quality, the stated goal of Title I of the Elementary and Secondary Education Act (ESEA) is to provide children with access to "high quality education," and the multidimensional index is intended to allow for "meaningful differentiation" between schools. While any analysis focused on school quality necessarily cannot be rooted in a clear-cut operationalization of school quality, we aim in this study to establish the degree to which variation *cannot* be explained by school quality—therefore characterizing the extent to which the ESSA-mandated meaningful differentiation of schools is meaningfully differentiating school by school quality or something else.

The remainder of this paper proceeds as follows. First, we provide a brief review of the literature related to school grades, how they are used, the extent to which they capture school quality versus other factors outside schools' locus of control, and the problem with proficiency as a measure of school quality. We turn next to the Florida context, including a brief history of the state's school accountability system and the way it measures school quality and assigns grades in the ESSA era. Next, we describe our data and methods, followed by our results. We conclude with a discussion of findings, implications for policy related to school grades, and directions for future research.

Literature Review

As of 2021, 11 states graded their public schools on A-F scales and another 19 used some other kind of rating system, either as a numeric index (14) or a star system (5) (Education Commission of the States, 2021). Florida, then, is among the majority of states in its reporting of school grades as part of its approach to school accountability under ESSA. Florida is also a bellwether state in school accountability—having begun assigning A-F grades in 1999 (Figlio & Loeb, 2011), before NCLB began requiring in-depth reporting of student achievement. Thus, Florida's school accountability system under ESSA provides a useful context through which to consider ESSA-compliant school quality measures.

In this literature review, we begin with a brief discussion of measurement, highlighting that different ways of measuring school quality will capture different factors—many of which are unrelated to school quality. We then discuss the “school quality” construct itself and the complications associated with establishing an agreed-upon measure of school quality. Next, we overview the evidence showing that families use school grades to make these decisions—underscoring the importance of these grades for informational purposes. We conclude by summarizing the literature showing that these decisions lead to greater segregation and reduced opportunity for already marginalized students.

When school accountability systems rate schools based on proficiency rates, they hold them accountable for their students' educational opportunities since birth, rather than the teaching and learning that actually occurs within the school itself (Harbatkin & Wolf, 2023; Heck, 2006; Kim & Sunderman, 2005; Krieg & Storer, 2006; Reardon, 2019). Research on letter grades from accountability systems before ESSA shows that the grades failed to meaningfully differentiate between schools after controlling for student and school characteristics (Adams, Forsyth, Ware, & Mwavita, 2016; Adams, Forsyth, Ware, Mwavita, Barnes et al., 2016). Recent research found that school-level measures continue under ESSA to be highly correlated with student demographics, though there is some variation by measure type (Atchison et al., 2023; Le Floch et al., 2023; Pivovarova & Powers, 2024). This is consistent with a large swath of prior research from the NCLB era showing that proficiency rates—which rely on arbitrary pass thresholds—fail to capture distributional shifts in achievement, can misrepresent longer-term trends, and are easily subject to gaming and strategic behavior (Balfanz et al., 2007; Ballou & Springer, 2016; Darling-Hammond, 2018; Heck, 2006; Ho, 2008; Reback, 2008). Indeed, there is evidence from NCLB that its emphasis on proficiency induced schools to target so-called “bubble students” around the proficiency threshold to the detriment of the lowest achieving students (Booher-Jennings, 2005). A theoretical benefit of ESSA is its move away from the proficiency focus, and some early research on ESSA school improvement has shown that it has successfully shifted the focus of improvement efforts to the students who stand to gain the most (Burns et al., 2023). However, little is known thus far about how ESSA's multidimensional measure may have facilitated differences in state reporting of school quality.

One challenge inherent in developing measures of school quality is there is no unequivocal definition of educational quality (Dijkstra et al., 2017; Schneider et al., 2017, 2018). Is a high-quality school (or teacher) one that increases test scores (Rivkin et al., 2005)? Fosters a positive, welcoming, and supportive school climate so that students want to attend (Gershenson, 2016; Hamlin, 2021; Jackson, 2012)? Achieves consistently high graduation rates or later college degree completion (Dynarski et al., 2013; Robertson et al., 2016)? Contributes to better longer-term outcomes, such as higher earnings and less criminal justice system engagement, of students who previously attended (Bernal et al., 2016; Chetty et al., 2011)? Produces graduates who become civic-minded citizens (Lenzi et al., 2014; Lin, 2015)? In the absence of an unambiguous measure of school quality, research aiming to examine the connection between school accountability ratings and school quality tend to define school quality not by what it is but by what it is not—in other words, true measures of school quality should reflect something other than the demographics and socioeconomics of the student populations served (Adams, Forsyth, et al., 2016; Harbatkin & Wolf, 2023; Ho, 2008; Hough et al., 2016; McEachin & Polikoff, 2012; Pivovarov & Powers, 2024; Reardon, 2019). One way to accomplish this is to decompose the variance in school grades into the part that is explained by demographic and socioeconomic factors outside the school’s control, and the part that is not explained. When a large share of the variation of a school grade is explained by out-of-school factors, there is little actual signal remaining that could reflect true school quality. In turn, the grades that states report reflect opportunity to learn outside of the school building rather than the learning that occurs inside the school building (Reardon, 2019).

Public reporting of school ratings matter because parents prefer effective schools (Denice & Gross, 2016; Rothstein, 2006), and given the opportunity, will select into the schools that they perceive to be most effective toward the aims that are important to them. Well-resourced families engage in Tiebout sorting, moving to neighborhoods that they perceive to have the highest quality schools (Bayer et al., 2004). Given information on school quality, parents are more likely to choose higher performing schools that are available to them (Hastings & Weinstein, 2008), which in turn, can result in changes in housing values, as parents pay more for houses in “better” school zones (Black, 1999; Figlio & Lucas, 2004; Rothstein, 2006). In contexts with robust school choice, people seek out school quality information online with greater intensity given expansions to accountability-driven school choice (Lovenheim & Walsh, 2018), more evidence that families leverage information such as school grades in school choice decisions. The way in which that information is reported may matter as well—there is evidence that changes to the format or order of information presented has the capacity to nudge parents to make different decisions for their children (Glazer et al., 2018; Schneider et al., 2018). Disseminating information on student growth in particular can induce parents to choose schools in a way that will contribute to desegregation efforts (Houston & Henig, 2023). If parents are making choices about schools based on student achievement alone, they may not be making the best choice for their child because the schools with the highest student achievement are not necessarily the most effective schools at raising test scores (Hough et al., 2016; Reardon, 2019) or non-test score outcomes (Beuermann et al., 2023). In sum, the grades that are posted—and the information contained in them—matter because parents use them to make choices about where to send their children to school.

As school choice has expanded nationwide, there has been growing concern that segregation would increase as families select into increasingly homogenous learning contexts (Frankenberg, 2018; Garcia, 2008; Kotok et al., 2017). This is particularly relevant in the context of school accountability because ESSA—along with its predecessor, NCLB—includes language that allows children zoned to schools identified as low performing to transfer to another district school, if available. In some cases, states may also choose to close or take over designated schools, leading to a loss of local control. Indeed, there is evidence that accountability systems can lead to school closures

in already underserved communities and can exacerbate segregation (Balfanz et al., 2007; Davis et al., 2015; Hasan & Kumar, 2019; Lee & Lubienski, 2017; Lipman, 2017). Thus, the design of ESSA accountability systems affects which schools are designated as low performing and therefore stand to lose students to other neighborhood schools as a result of the designation. In order to fill this gap about how ESSA-mandated multidimensional measures may have facilitated differences in state reporting of school quality, we examine whether Florida's ESSA-era school grading system appears to be capturing—and reporting out—measures of school quality that are not as clearly confounded by school and community demographics that plagued NCLB-era systems.

School Accountability in Florida

Florida's multidimensional school rating index includes 11 components—four proficiency rates (ELA, math, science, and social studies), four learning gains measures (ELA and math, respectively, overall and for the lowest achieving 25% of students), graduation rate, and “acceleration,” which captures advanced coursetaking, dual enrollment, and industry certification (Florida Department of Education, 2021). Broadly, learning gains are initially calculated as a dichotomous measure at the student level, where students count as having made sufficient gains in a particular subject if they increase by one achievement level or sublevel or remain at the same proficient-or-above achievement level but increase their scale score.¹ This measure is then aggregated to the school level for math and ELA, respectively, overall and then for the lowest achieving quartile of students in each subject based on prior years. It is reported as the percent of students making learning gains. Schools are then assigned letter grades based on the percentage of total points earned, including proficiency rates, learning gains, graduation rates, and acceleration.

The components that make up these letter grades are those in the state's meaningful differentiation index under ESSA, with an additional indicator focusing on the progress of English learners, in accordance with ESSA requirements. The state places schools in Comprehensive Support and Improvement (CSI) and Targeted Support and Improvement (TSI) under ESSA based on a combination of the federal index and school grades. Schools are designated as CSI if they either have a current grade of D or F, have a graduation rate of 67% or lower, have an overall federal percent-of-points index (which includes the letter grade components plus EL progress) of 40% or lower, or are a TSI school with a subgroup federal percent-of-points index 40% or lower for six years. Schools can exit CSI status when they reach the points index threshold in a subsequent year, though those receiving an “F” grade cannot exit until they implement a two-year turnaround plan. Previously-F schools that do not earn a “C” grade or higher after two years must close or turn over operations to a charter or an external operator. Schools are designated as TSI if any subgroup's performance on the federal percent of points index is 31% or lower over three years, or if any subgroup's performance on the federal percent of points index is 40% or lower in the current year. They can exit TSI status once they improve subgroup performance to 41% or higher on the federal percent of points index. If no improvement has been made within six years, the school moves to CSI status (Florida Department of Education, 2018).

These findings further suggest that accountability measures may not always accurately reflect the effectiveness of schools in serving their students, particularly those from disadvantaged backgrounds, highlighting the need for more nuanced and equitable assessment frameworks.

¹ For more on Florida's highly specific criteria for meeting gains, please see the state's school grades calculation guide, beginning on page 13:

<https://www.fldoe.org/core/fileparse.php/18534/urlt/SchoolGradesCalcGuide21.pdf>

During to the COVID-19 pandemic, school districts and charter school governing boards were granted the flexibility to opt out of reporting their school grade and/or school improvement rating. To be eligible for a school grade, a school needed to have tested 90% or more of its eligible students during the 2020-21 academic year, which is less than the usual 95%. Schools that did not opt in or that failed to meet eligibility criteria did not receive a school grade or school improvement rating for the 2020-21 school year. School accountability resumed in full for the 2021-2022 school year, and CSI/TSI school identification resumed in fall 2022 (Florida Department of Education, 2021).

Data, Sample, and Methods

Data and Sample

Guided by previous studies addressing accountability and school grades (Adams, Forsyth, Ware, & Mwavita, 2016; Adams, Forysth, Ware, Mwavita, Barnes et al., 2016; Pivovarova & Powers, 2024) we used available variables from publicly available school- and county-level data from four sources to answer our research questions. Our outcomes of interest, along with school demographic data on race and free or reduced lunch eligibility, school level (i.e., elementary, middle, high, middle/high), and governance model (i.e., traditional public school or charter) come from the Florida Department of Education (FLDOE) Education Data Archive for the 2021-22 school year. We use school-level data on students with disabilities from the National Center for Education Statistics (NCES) Common Core of Data (CCD) through the Urban Institute's Education Data Portal. We draw on U.S. Census American Community Survey (ACS) one- and five-year estimates on county race/ethnicity, educational attainment, Temporary Assistance for Needy Families (TANF) and Supplemental Nutrition Assistance Program (SNAP) eligibility, and poverty. Finally, we draw on annual county-level unemployment rate data from the U.S. Bureau of Labor Statistics (BLS).

We excluded from our sample 25 virtual schools and 16 schools without complete data. In total, we have about 3,400 schools with school accountability grades in 2022 in all 67 counties. Nearly two-thirds of schools are elementary, with the remainder split between middle and high schools (with a small subset of 110 combination schools). More than 80% are traditional public schools (TPS), and less than 20% are charters. Fifty-eight percent are located in suburbs or towns, 29% in urban settings, and 14% in rural areas.

Outcomes

We draw seven outcome measures from the FLDOE school report card. These are the A–F letter grade assigned by the state system, the 0-100 state percent-of-points index, as well as proficiency, learning gains, and learning gains of the bottom 25% in math and ELA, respectively. The A-F grade is based on the sum of points earned for each component in the state index system, using the following percentages: A: 62% of points or greater, B: 54% to 61% of points, C: 41% to 53% of points, D: 32% to 40% of points, or F: 31% of points or less. We code these grades on a 0-4 scale grade-point average scale, with 0 representing a grade of F, 1 for D, 2 for C, 3 for B, and 4 representing an A grade. Each of the other measures is a percentage.

Table 1, Panel A, provides summary statistics for each of these outcomes, overall and then by school level and governance structure. The average school in our sample has a grade of about 2.8, or a C+, and earned 454 of an average of 802 total possible points (while there are 1,100 total points possible, most schools did not meet minimum inclusion thresholds for all 11 components and the modal school's grade was based on seven components). The average school had about a 50% proficiency rate, and met nearly 60% of math and ELA learning gains, respectively. Learning gains of the lowest achieving 25% was slightly lower, at about 48% in ELA and 55% in math. Figures

were relatively similar across school levels and governance structure, with minor exceptions. Elementary schools had higher achievement and gains than other school levels. High schools earned more points on average because they include the graduation rate component, which is worth up to 100 points, and the others do not. Charters had higher proficiency rates but only marginally higher gains than traditional public schools.

Table 1

Summary Statistics Overall, by School Level, and by Governance Structure

	Overall	Elementary	Middle	High	TPS	Charter
<i>Panel A. Outcomes</i>						
School Grade	2.8 (1.0)	2.8 (1.0)	2.7 (0.9)	2.9 (0.9)	2.8 (1.0)	3.0 (1.0)
Total Points Earned	454.4 (124.8)	411.3 (100.9)	489.3 (97.7)	567.6 (128.4)	446.2 (118.9)	493.5 (143.5)
ELA Achievement	53.4 (16.9)	54.6 (16.8)	49.1 (16.2)	52.6 (17.2)	52.1 (16.5)	59.5 (17.5)
ELA Learning Gains	56.5 (11.2)	60.1 (10.1)	48.1 (9.1)	52.0 (10.6)	56.2 (11.1)	58.3 (11.1)
ELA Learning Gains of the Lowest 25%	47.9 (12.6)	51.7 (11.6)	39.0 (10.1)	42.8 (12.6)	47.4 (12.5)	50.7 (13.0)
Mathematics Achievement	54.4 (18.1)	57.7 (16.8)	52.0 (17.9)	44.7 (18.9)	53.7 (17.7)	57.7 (19.7)
Mathematics Learning Gains	60.2 (13.4)	63.8 (12.4)	57.4 (11.3)	49.8 (12.9)	60.0 (13.3)	61.2 (14.2)
Mathematics Learning Gains of the Lowest 25%	54.8 (13.5)	56.1 (14.1)	54.9 (10.8)	49.9 (12.6)	54.2 (13.1)	57.7 (14.8)
<i>Panel B. School-level variables</i>						
Economically disadvantaged	66.3 (28.4)	69.1 (28.7)	67.4 (25.8)	58.5 (26.1)	69.1 (26.9)	52.8 (31.1)
Black	22.5 (21.9)	23.1 (22.7)	22.8 (21.1)	20.4 (19.7)	23.1 (22.1)	19.4 (20.4)
White	33.9 (24.3)	32.7 (24.3)	33.9 (24.1)	36.6 (24.5)	35.0 (24.3)	28.7 (23.6)

	Overall	Elementary	Middle	High	TPS	Charter
Hispanic	33.7 (24.0)	33.6 (24.0)	34.9 (24.1)	34.9 (24.6)	32.2 (22.9)	41.0 (27.8)
Asian or Pacific Islander	4.2 (3.5)	4.4 (3.5)	3.5 (3.4)	3.7 (3.5)	4.1 (3.3)	4.7 (4.3)
Indigenous	0.8 (1.3)	0.8 (1.5)	0.8 (1.0)	0.7 (0.8)	0.8 (1.0)	0.9 (2.4)
2+ races	5.0 (3.0)	5.3 (2.9)	4.1 (3.0)	3.7 (2.2)	4.9 (2.8)	5.3 (4.1)
Students with Disabilities	16.5 (6.1)	16.5 (6.0)	18.5 (5.7)	14.9 (5.5)	17.4 (5.6)	11.5 (5.9)
English Learners	11.8 (11.8)	14.5 (13.2)	8.2 (7.0)	6.0 (5.8)	11.9 (11.9)	11.3 (11.5)
Out-of-Field Teachers	9.7 (9.7)	9.2 (9.8)	11.7 (9.5)	9.5 (8.8)	8.6 (7.8)	15.8 (15.1)
Teachers Rated Highly Effective	67.3 (28.2)	68.2 (28.2)	64.7 (27.9)	67.8 (26.6)	69.6 (26.0)	52.9 (36.4)
Teachers Rated Unsatisfactory	0.2 (1.3)	0.2 (1.1)	0.3 (1.6)	0.2 (1.5)	0.2 (1.2)	0.3 (1.3)
<i>Panel C. County-level variables</i>						
Unemployment Rate	4.7 (0.7)	4.7 (0.6)	4.7 (0.7)	4.7 (0.7)	4.6 (0.7)	4.8 (0.7)
SNAP	14.8 (5.2)	14.8 (5.1)	14.6 (5.1)	15.3 (5.6)	14.6 (5.1)	15.9 (5.6)
Child Poverty	20.3 (2.4)	20.4 (2.4)	20.3 (2.4)	20.3 (2.5)	20.3 (2.5)	20.4 (2.3)
Percent above Bachelor's	31.3 (6.9)	31.5 (6.7)	31.6 (6.7)	30.6 (7.4)	31.2 (7.2)	31.9 (5.6)
Black	14.8 (7.4)	15.0 (7.3)	14.7 (7.3)	14.5 (7.3)	14.6 (7.3)	15.9 (7.4)
White	50.4 (20.3)	50.1 (19.9)	50.8 (20.2)	49.9 (21.7)	51.8 (19.9)	43.9 (21.0)

	Overall	Elementary	Middle	High	TPS	Charter
Hispanic	27.2 (19.2)	27.3 (18.9)	27.0 (18.9)	28.4 (20.8)	26.0 (18.5)	33.0 (21.3)
Indigenous	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)
Asian or Pacific Islander	3.0 (1.5)	3.1 (1.5)	3.0 (1.5)	2.8 (1.5)	3.0 (1.5)	2.9 (1.4)
2+ races	3.7 (1.1)	3.7 (1.1)	3.7 (1.1)	3.6 (1.1)	3.7 (1.1)	3.4 (1.1)
N	3,406	2,150	572	574	2,812	594

Note: Includes all schools with accountability grades in Florida in 2021-22. Sum of elementary, middle, and high schools does not add up to total *N* because of 110 combination middle/high schools reflected as part of first column but not included as their own column for simplicity.

Predictors

School-level Covariates. From the FLDOE, we draw the share of students eligible for free or reduced lunch within a school,² the share of students by race and ethnicity, and the share of English learners (ELs). Because the state suppresses economic disadvantage data for schools with fewer than 10 students who qualify as economically disadvantaged, we impute suppressed values with 0.05%. Race and ethnicity categories reported by FLDOE include White, Black, Hispanic, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native, and two or more races. Due to small samples, we combine the Asian and Native Hawaiian categories into a single category. For both race/ethnicity and ELs, the state suppressed values for schools that had more than zero but fewer than 10 students in a group. In our analyses, we imputed these suppressed values with the midpoint of five students. School disability rates, drawn from the CCD, are operationalized as the percentage of students with disabilities served under Section 504 and students with disabilities served under IDEA.

For supplementary analyses using variables that do represent measures of school quality, we also drew measures from FDOE on teacher effectiveness and classes taught by out-of-field teachers. We focus on teacher qualifications because they are measurable and because they are the most important input in a child's education (e.g., Chetty et al., 2014). Specifically, we use the share of teachers rated as ineffective on Florida's teacher evaluation system. Florida counts a course as having an out-of-field teacher if the primary instructor does not have the qualifications required for that course and subject. Teacher effectiveness is determined by the state's teacher evaluation system, which provides teachers with a rating of highly effective, effective, needs improvement/developing, or unsatisfactory. Because the vast majority of teachers are rated effective or highly effective, we focus in this analysis on the upper and lower ends of the rating scale with the shares rated unsatisfactory and highly effective, respectively. Finally, we draw the locale code of the school's

² In Florida, a student qualifies for free or reduced-price meals either through Direct Certification determination or by extension of eligibility to the household. A student can also be eligible for free meals solely based on eligibility survey results. The eligibility for free or reduced-price meals is measured consistently across TPS and charter schools.

physical address from the U.S. Census, and collapse the Census locale codes into three categories—urban, suburban/town, and rural.

As shown above in Table 1, Panel B, the average school in our sample is about 70% economically disadvantaged, as measured by eligibility for free- or reduced-price lunch and including the community eligibility provision, 23% Black, 34% Hispanic. TPSs have substantially higher poverty than charters, at nearly 70% compared with 53% for charters. TPSs serve more White students and students with disabilities, respectively, and charters serve more Hispanic students. About 10% of classes are taught by out-of-field teachers, though this figure is much higher in charters (16%) than TPSs (9%). Across all subgroups, most teachers are rated highly effective and less than half a percent are rated unsatisfactory. Teachers in charters are least likely to receive the highly effective rating, with just over half as highly effective compared with about two-thirds in each of the other subgroups.

County-level Covariates. To examine the extent to which community characteristics are associated with school grades, we merge school accountability and demographic data with county-level data based on the school's physical location. From the ACS, we include five-year county estimates of Black, Hispanic, Asian or Pacific Islander, Indigenous, and two or more races, with White as the reference category. We also draw educational attainment data from the ACS using 5-year estimates to construct a variable for percent of county residents with a bachelor's degree or above. Because school-level economic disadvantage is a blunt measure of poverty (Hashim et al., 2023; Owens et al., 2016), we also draw on two more nuanced measures of county-level socioeconomic status that are especially relevant to families with children—SNAP eligibility and child poverty. SNAP eligibility represents the ACS one-year estimated percent of households in a county that were eligible for SNAP within the past 12 months for 2021. Child poverty is calculated as the estimated percent of under-18 population in poverty in the ACS 2022 five-year averages. For unemployment rate, we use the 2021 mean county-level unemployment rate from the BLS.

The average school in our sample was located in a county with an unemployment rate of about 4.7%, SNAP rate of about 15%, and child poverty rate of about 20%. About three in 10 residents had bachelor's degrees, half were White, 27% Hispanic, and 15% Black. Charters were located in counties home to more Hispanic and Black residents and fewer White residents than TPS schools.

The included covariates emerge from an existing literature that has established a meaningful association between these demographics, socioeconomics, and existing measures of school quality (Adams, Forsyth, Ware, & Mwavita, 2016; Adams, Forsyth, Ware, Mwavita, Barnes et al.; Harbatkin & Wolf, 2023; Hough et al., 2016; McEachin & Polikoff, 2012; Reardon, 2019). They do not capture a comprehensive set of out-of-school factors that may confound school grades, but they do provide a reasonable starting point for any policymaker aiming to test the extent to which grades stemming from a proposed accountability system may be confounded by out-of-school factors because they are widely measured and accessible for all schools.

Methods

We answer our research questions using a combination of descriptive statistics and descriptive regressions. We predict each of the eight outcomes (i.e., A-F school grade, percent of total points, and ELA and math proficiency, gains, and gains of bottom 25%) as a function of school covariates, then as a function of county covariates, and finally as a function of both school and county covariates. The initial, school-covariate-only model predicting school grade, takes the form

$$Grade_s = \beta_0 + \sigma \mathbf{X}'_s + \pi_s + \delta_s + \mu_s + \varepsilon_s \quad (1)$$

predicting the school grade for school s . \mathbf{X}' is a vector of school-level covariates including economically disadvantaged percent, school race/ethnicity percentages with White as the omitted reference category, percent of students with disabilities, and percent of ELs; π represents school-level fixed effects (elementary, middle, high, middle/high), δ represents charter fixed effects, μ represents locale fixed effects (urban, rural, suburban/town) and ε is an idiosyncratic error term.

We then run a parallel model that replaces the vector of school characteristics with a vector of county-level characteristics, \mathbf{Y}' , that includes county race/ethnicity percentages, unemployment rate, SNAP, child poverty, and county residents with a bachelor's degree or above:

$$Grade_s = \beta_0 + \gamma \mathbf{Y}'_s + \pi_s + \delta_s + \mu_s + \varepsilon_s \quad (2)$$

We then estimate a including both school and county-level covariates, taking the form

$$Grade_s = \beta_0 + \sigma \mathbf{X}'_s + \gamma \mathbf{Y}'_s + \pi_s + \delta_s + \mu_s + \varepsilon_s \quad (3)$$

We repeat the same set of models for each of the eight outcomes, and then compare the adjusted R^2 for each of the different outcomes to quantify the extent to which each outcome is explained by school- and county-level sociodemographic variables that are unrelated to school effectiveness. An adjusted R^2 closer to one for a given outcome provides evidence that the school accountability score is driven more by out-of-school factors than school quality, while a value closer to zero implies that a given school accountability score is driven less by these factors—at least observed factors—and therefore may better capture in-school factors. We also run these same models separately for TPS only and charter only to examine whether there are differences by governance structure.

One important limitation is that some of these socioeconomic and demographic variables may in fact be associated with true differences in school quality. For example, high quality teachers tend to sort to higher socioeconomic status schools (Ingersoll, 2004; Jackson, 2009). This can be interpreted as either an alternative explanation or a mechanism. For example, experienced and highly effective teachers may select out of less advantaged schools due to potentially malleable factors such as working conditions, but it is also the case that working conditions tend to be more challenging when schools lack adequate resources to for their teachers and students (Harbatkin, Nguyen, et al., 2023; Ingersoll, 2001; Redding & Nguyen, 2020). Ultimately, it is possible that more disadvantaged schools may provide lower quality education, on average, than their more advantaged counterparts because they face greater challenges recruiting experienced and highly effective teachers (Engel et al., 2014). Thus, to examine the extent to which our out-of-school factors may be confounded by true measures of school quality and may therefore lead our initial model to overstate their *direct* contributions to school grades, we run a partial mediation model adding variables representing out-of-field teaching, highly effective teachers, and unsatisfactory teachers. Following Baron and Kenny (1986), we run this model in two steps. First, we replace the outcome in Equations 1 through 3 with each teacher quality variable, respectively. This provides us with the estimate on Baron and Kenny's path "a", which is the estimated relationship between the school and teacher covariates (predictors) and teacher quality (mediator). The estimates on the school and county variables provide their estimated relationship with teacher quality. To the extent that these are significant, teacher quality may mediate the relationship between our out-of-school factors and school grades. Next, we add the teacher quality variables to the right side of Equation 3, with the model taking the form:

$$Grade_s = \beta_0 + \sigma X'_s + \gamma Y'_s + \eta Z'_s + \pi_s + \delta_s + \mu_s + \varepsilon_s \quad (4)$$

where Z'_s is a vector of the three school-level teacher quality variables and the rest of the model remains the same. To the extent that the estimates on our out-of-school factors are attenuated from Equation 3 to Equation 4, we can assume that they are correlated with true measures of school quality.³ We can then calculate the share of the Equation 3 coefficient estimates that are explained by teacher quality differences and therefore may overestimate the contribution of out-of-school factors to the school grade measures. We do this through simple division: For example, if a coefficient estimate A is 0.30 in Equation 3, and then attenuates to 0.20 in Equation 4, then we can conclude that one-third of the relationship between variable A and the school grade outcome can be explained by the teacher quality variables i.e., $\left(\frac{0.30-0.20}{0.30} = 0.33\right)$.

Finally, we can decompose the covariance between the school grade outcome and out-of-school factors by running one final model, predicting the school grade outcome as a function of *only* the teacher quality variables and examine the adjusted R^2 from that model:

$$Grade_s = \beta_0 + \eta Z'_s + \varepsilon_s \quad (5)$$

We then compare the adjusted R^2 from Equations 3, 4, and 5 as:

$$AdjR^2(Eq5) - [AdjR^2(Eq4) - AdjR^2(Eq3)], \quad (6)$$

which is bounded between zero and one. If the teacher quality variables were completely uncorrelated with our observed out-of-school factors, the difference in adjusted R^2 between Equation 3 and Equation 4 would match the adjusted R^2 in Equation 5, and Equation 6 would equal zero. If 100% of the covariance between the teacher quality variables and the school grade also covaried with the observed out-of-school factors, then Equation 6 would equal one. The solution to Equation 6 therefore allows us to calculate exactly what share of the original adjusted R^2 is confounded by teacher quality and therefore may reflect true school quality.

There are several limitations to these analyses and the resulting interpretations. Our analyses are purely descriptive and cannot reasonably capture all out-of-school factors that contribute to school quality. As described above, it is also the case that the model excludes true measures of school quality that are associated with the measured out-of-school factors. Our assumption is the contribution of the first limitation outweighs the second; in other words, we believe our model fit estimates understate the contribution of out-of-school factors to school ratings. However, to the extent that unobserved factors contribute to school ratings consistently across outcome ratings, our comparisons *across* ratings will not be confounded by these unobserved factors even if the R^2 s underestimate the total amount of variation that can be explained. An additional limitation stems from the use of a single year of post-pandemic data that may or may not be generalizable moving forward given the pandemic's real effects on various aspects of children's learning. The widespread

³ This assumption holds when the sample is the same across the two models being compared, as the only difference between the models is the presence of the teacher quality variables. Thus, in order to make the comparison, we rerun Equation 3 on the sample of 2,971 schools for which we have complete teacher quality variables (i.e., we drop 102 schools from the full models that did not report school-level teacher quality measures) and compare coefficients across models run on the same sample. Estimates without those 102 schools were substantively similar and are provided in a supplemental appendix (Appendix Table A4).

disruptions caused by the pandemic introduced novel factors to the learning environment including effects on student well-being, remote learning, and disparities in access to technology (Finch & Hernández Finch, 2020; Harbatkin, Strunk, et al., 2023). The contribution of these disruptions will be heavily weighted in an analysis of a single year of post-pandemic data. To the extent that schools are able to effectively mitigate learning disruptions, the relationship between out-of-school factors and school quality measures may attenuate in future years. On the other hand, if achievement gaps grow due to inequitable opportunity to learn, the relationship may grow stronger.

Additionally, findings based on a single year of data are subject to idiosyncratic year-to-year variation, though we believe that they provide an informative post-pandemic snapshot of the association between sociodemographic factors and school grades. To the extent that this variation is atypical, the analyses could either over- or under-state the magnitude of the coefficient estimates and model fit. However, we feel that the urgency of understanding the role of these factors immediately post-pandemic—especially given that the federal Elementary and Secondary Education Act (ESEA) that underlies the ESSA index is currently overdue for reauthorization (DeBray et al., 2022)—outweighs the benefits of additional years of post-pandemic data. Finally, Florida is just one state with a unique demographic and political context that may contribute to the generalizability of our findings. However, as a national leader in school accountability policy, Florida has long been a lodestar for other state policies, and there is reason to believe that our findings will have broader applicability to state accountability systems nationally.

Findings

Main Findings

Table 2 displays our results for RQ1, showing the results of Equation 3 predicting each of the separate components of the letter grade separately, with proficiency, learning gains, and learning gains of the bottom 25%, respectively, in rows 1-3 for ELA and 4-6 for math. School-level covariates are at the top, followed by county-level covariates.⁴ Because all predictors are scaled as percentages (0-100), it is possible to compare the magnitude of coefficient estimates. There are four takeaways from these coefficient estimates. First, the school-level variables that are expected to predict proficiency do so, and in the expected direction. For example, a 1 percentage point increase in economic disadvantage is associated with a 0.22 point decrease in ELA proficiency rate (Column 1) and a 0.19 decrease in math proficiency rate (Column 4). That means a one-standard deviation increase in economic disadvantage is associated with a 6.5 percentage point decrease in ELA proficiency and a 5.4 percentage point decrease in math proficiency, holding all other variables constant (the magnitude is larger in the models with just school covariates). The relationship between proficiency and Black student percentage is even larger, with a 1pp increase in Black students being associated with a 0.39 point decrease in ELA proficiency and 0.42 point decrease in math proficiency. Put another way, a one-standard deviation increase in Black students is associated with an 8.6 percentage point decrease in ELA proficiency and a 9.2 percentage point decrease in math proficiency.

Second, demographic variables are predictive of each of the six measures but are generally most predictive of proficiency and least predictive of learning gains of the bottom 25%. For example, economic disadvantage is about two and a half times more predictive of ELA proficiency than ELA learning gains, and four times more predictive of ELA proficiency than ELA learning gains of the bottom 25%. Black percentage is about 2.2 times more predictive of ELA proficiency

⁴ Regression results for Equations 1 and 2, predicting each as a function of school and county covariates, respectively, on their own, are provided in Appendix Tables A1 (school only) and A2 (county only).

than learning gains and nearly 10 times more predictive of ELA proficiency than learning gains of the bottom 25%. In math, the coefficient on Black percentage predicting learning gains of the bottom 25% is attenuated to the point of statistical insignificance. Similarly, Hispanic and EL percentage, respectively, are more predictive for proficiency than learning gains, and then attenuates to statistical insignificance in both math and ELA learning gains of the bottom 25%.

Third, as unemployment rate increases, proficiency and learning gains decrease. In fact, unlike in the case of the school-level variables, learning gains appear to be *more* responsive than proficiency to shifts in unemployment rate. This is also the case for county-level child poverty rate.⁵ Finally, after controlling for school-level variables, county-level variation in race and ethnicity does not consistently move in the expected direction. This is likely because of within-county sorting and segregation.

Table 2

Regressions predicting each subcomponent as a function of school and county covariates

	ELA			Math		
	(1)	(2)	(3)	(4)	(5)	(6)
	Proficiency	Learning gains	Bottom 25% learning gains	Proficiency	Learning gains	Bottom 25% learning gains
<i>School-level variables</i>						
Economic disadvantage	-0.224*** (0.009)	-0.087*** (0.008)	-0.056*** (0.011)	-0.188*** (0.012)	-0.088*** (0.011)	-0.071*** (0.013)
Black	-0.394*** (0.014)	-0.177*** (0.012)	-0.040* (0.017)	-0.423*** (0.018)	-0.155*** (0.017)	-0.006 (0.020)
Hispanic	-0.109*** (0.017)	-0.082*** (0.014)	-0.023 (0.020)	-0.162*** (0.021)	-0.081*** (0.019)	-0.013 (0.022)
Asian or Pacific Islander	0.650*** (0.055)	0.412*** (0.046)	0.323*** (0.065)	0.497*** (0.069)	0.267*** (0.064)	0.101 (0.074)
Indigenous	-0.037 (0.129)	0.115 (0.109)	0.214 (0.153)	0.161 (0.163)	0.107 (0.150)	0.436* (0.176)
2+ races	-0.225** (0.085)	-0.057 (0.072)	0.225* (0.105)	-0.417*** (0.108)	-0.051 (0.099)	-0.027 (0.120)
Students with Disabilities	-0.339*** (0.034)	-0.191*** (0.029)	-0.124** (0.041)	-0.398*** (0.043)	-0.194*** (0.040)	-0.084 (0.047)
English learners	-0.345*** (0.021)	-0.049** (0.018)	-0.047 (0.025)	-0.257*** (0.027)	-0.060* (0.025)	-0.023 (0.029)

⁵ This is shown more clearly in the supplemental table (online appendix Table A2) with county only variables.

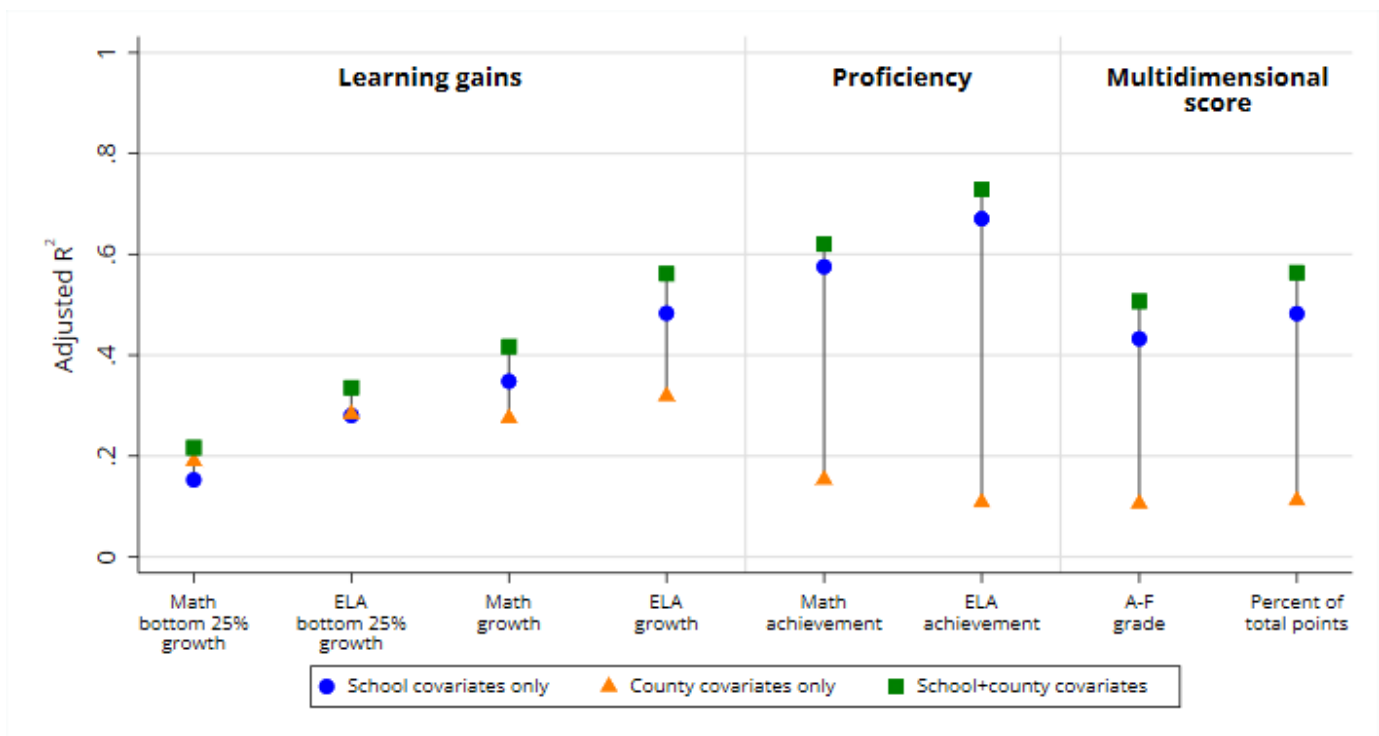
	ELA			Math		
	(1) Proficiency	(2) Learning gains	(3) Bottom 25% learning gains	(4) Proficiency	(5) Learning gains	(6) Bottom 25% learning gains
<i>County-level variables</i>						
Unemployment Rate	-1.582** (0.522)	-1.455** (0.442)	-2.145*** (0.616)	-5.028*** (0.661)	-4.943*** (0.609)	-4.875*** (0.708)
SNAP Eligibility	0.093 (0.091)	-0.076 (0.077)	0.239* (0.108)	0.373** (0.116)	-0.043 (0.107)	0.084 (0.124)
Child poverty	0.066 (0.117)	-0.129 (0.099)	-0.246 (0.137)	0.077 (0.148)	-0.225 (0.136)	-0.284 (0.158)
Above Bachelor's	0.455*** (0.059)	0.264*** (0.050)	0.203** (0.069)	0.280*** (0.074)	0.051 (0.068)	-0.031 (0.079)
Black	0.420*** (0.040)	0.384*** (0.034)	0.245*** (0.047)	0.332*** (0.050)	0.445*** (0.046)	0.408*** (0.054)
Hispanic	0.212*** (0.034)	0.250*** (0.029)	0.174*** (0.040)	0.129** (0.043)	0.349*** (0.040)	0.330*** (0.046)
Indigenous	-8.091*** (2.368)	-4.286* (2.006)	-4.559 (2.786)	-8.217** (3.001)	-3.806 (2.762)	-3.695 (3.202)
Asian or Pacific Islander	-3.018*** (0.242)	-1.757*** (0.205)	-1.444*** (0.285)	-1.505*** (0.307)	-0.733** (0.282)	-0.749* (0.328)
2+ races	0.775* (0.363)	0.557 (0.308)	0.129 (0.429)	-0.990* (0.460)	0.178 (0.424)	0.234 (0.493)
Constant	78.721*** (4.533)	69.749*** (3.841)	59.017*** (5.343)	106.110*** (5.745)	92.558*** (5.288)	78.424*** (6.145)
<i>N</i>	3,073	3,073	3,047	3,073	3,073	3,047
<i>R</i> ²	0.731	0.565	0.340	0.623	0.421	0.222
Adjusted <i>R</i> ²	0.729	0.562	0.335	0.620	0.417	0.216

Note: Estimates from regressions predicting each subcomponent as a function of school and county covariates. All models include school-level fixed effects, a charter indicator, and locale fixed effects. All outcomes measured on a 0-100 percentage scale. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To allow for a clear comparison of the extent to which each school quality measure is confounded by out-of-school factors, the adjusted R^2 for each model predicting each outcome is shown visually in Figure 1. The models predicting learning gains are shown in the first panel, followed by the models predicting proficiency, and then the multidimensional scores. The blue circles denote models including school covariates only (Equation 1), the orange triangles denote models including county covariates only (Equation 2), and finally the green squares represent Equation 3 that includes both school and county covariates. Here, it is clear that school and county covariates are highly predictive of proficiency rates, consistent with prior literature. Specifically, county covariates on their own explain 11% of ELA and 15% of math proficiency, while school covariates on their own explain 67% of ELA and 58% of math proficiency, and the combined models explain 73% of ELA and 62% of math proficiency. By comparison, the combined models explain only 56% of ELA and 42% of math gains, and 34% of ELA and 22% of math gains of the bottom 25%. In sum, each of these accountability system components are explained to some extent by factors outside of the school—but the contribution of external factors is strongest for proficiency and weakest for learning gains. Additionally, in alignment with a very large literature that consistently finds larger intervention effects for math than ELA, we show here that out-of-school factors are more predictive in ELA than in math.

Figure 1

Adjusted R^2 by Included Predictors and Model



To answer RQ2, Table 3 provides the results from Equation 3 predicting school grade and percent possible points. Columns 1 and 2 show results from the models predicting A-F grade (on a GPA scale of 0-4) and percent possible points (0-100%) for all schools, 3 and 4 show results from the same models for just TPS schools, and 5 and 6 show results from the same models just for charter schools.

There are three takeaways from the overall models. First, it is clear that the multidimensional measure of school quality is capturing out-of-school factors in addition to in-school factors. As shown by the adjusted R^2 s at the bottom of the table and the right-most panel of Figure 1 above, about half of the variation in the multidimensional A-F grade and 56% of the percent possible points measure can be explained by school and county characteristics. As expected, given that the measures contain each of the subcomponents, the adjusted R^2 s here are a weighted average of the subcomponents, with a magnitude lower than the proficiency measures and higher than the gains measures. Second, student characteristics, such as economic disadvantage and race, remain highly predictive of these grades. For example, the coefficient estimate of -0.012 on economic disadvantage suggests that a one standard deviation increase in economic disadvantage percent (28.4%) would take the average school from a 2.8 (C+) grade to just under 2.5 (C), holding all other school and county covariates constant. A one standard deviation increase in Black students (21.9%) would take the average school to about a 2.4 (C) grade. These translate to a loss of about 4 and 4.7 percentage points, respectively, on the percent of points index (Column 2). Third, there are two county-level covariates that are consistently meaningfully associated with school grades even over and above school covariates. Higher unemployment rate is associated with decreased school grades, and higher educational attainment is associated with increased school grades, over and above school covariates. For example, a 1 percentage point increase in unemployment rate would take the average 2.8/C+ school down to about 2.5/C, while a 10 percentage point increase in the percent of county residents with a bachelor's degree would take the average school up to nearly 3.0/B.

Table 3

Regressions predicting multidimensional school rating as a function of school and county covariates, overall and by governance structure

	Overall		TPS		Charter	
	(1)	(2)	(3)	(4)	(5)	(6)
	A-F grade	% possible points	A-F grade	% possible points	A-F grade	% possible points
<i>School-level variables</i>						
Economic disadvantage	-0.012*** (0.001)	-0.138*** (0.008)	-0.014*** (0.001)	-0.163*** (0.009)	-0.005** (0.002)	-0.067** (0.022)
Black	-0.019*** (0.001)	-0.215*** (0.012)	-0.017*** (0.001)	-0.194*** (0.013)	-0.026*** (0.004)	-0.275*** (0.042)
Hispanic	-0.009*** (0.001)	-0.082*** (0.014)	-0.007*** (0.001)	-0.072*** (0.015)	-0.011** (0.004)	-0.075 (0.044)
Asian or Pacific Islander	0.019*** (0.004)	0.390*** (0.047)	0.022*** (0.005)	0.383*** (0.050)	0.007 (0.011)	0.377** (0.131)
Indigenous	0.000 (0.010)	0.151 (0.110)	-0.027 (0.014)	-0.095 (0.155)	0.009 (0.027)	0.131 (0.327)
2+ races	-0.021** (0.007)	-0.187** (0.072)	-0.018* (0.008)	-0.123 (0.083)	-0.027 (0.014)	-0.330 (0.171)

	Overall		TPS		Charter	
	(1) A-F grade	(2) % possible points	(3) A-F grade	(4) % possible points	(5) A-F grade	(6) % possible points
Students with Disabilities	-0.012*** (0.003)	-0.233*** (0.029)	-0.008** (0.003)	-0.193*** (0.029)	-0.042*** (0.010)	-0.507*** (0.115)
English learners	-0.010*** (0.002)	-0.158*** (0.018)	-0.009*** (0.002)	-0.141*** (0.019)	-0.018*** (0.005)	-0.244*** (0.058)
<i>County-level variables</i>						
Unemployment Rate	-0.285*** (0.040)	-3.374*** (0.444)	-0.308*** (0.041)	-3.547*** (0.443)	0.066 (0.147)	-0.066 (1.754)
SNAP Eligibility	0.017* (0.007)	0.216** (0.078)	0.025*** (0.007)	0.274*** (0.078)	-0.017 (0.025)	-0.057 (0.294)
Child poverty	-0.011 (0.009)	-0.137 (0.099)	-0.010 (0.009)	-0.117 (0.097)	-0.003 (0.041)	-0.132 (0.494)
Above Bachelor's	0.016*** (0.005)	0.193*** (0.050)	0.012** (0.005)	0.155** (0.049)	0.053** (0.019)	0.559* (0.229)
Black	0.027*** (0.003)	0.315*** (0.034)	0.025*** (0.003)	0.301*** (0.034)	0.027* (0.010)	0.283* (0.124)
Hispanic	0.018*** (0.003)	0.202*** (0.029)	0.015*** (0.003)	0.173*** (0.029)	0.023* (0.010)	0.247* (0.119)
Indigenous	-0.543** (0.183)	-7.167*** (2.015)	-0.651*** (0.181)	-8.288*** (1.938)	1.085 (1.027)	13.124 (12.285)
Asian or Pacific Islander	-0.103*** (0.019)	-1.352*** (0.206)	-0.085*** (0.019)	-1.126*** (0.203)	-0.289*** (0.080)	-3.280*** (0.959)
2+ races	0.023 (0.028)	0.124 (0.309)	-0.003 (0.028)	-0.243 (0.304)	0.314* (0.124)	3.260* (1.483)
Constant	5.038*** (0.351)	84.100*** (3.856)	5.282*** (0.358)	86.915*** (3.840)	1.772 (1.361)	49.451** (16.277)
N	3,073	3,073	2,624	2,624	449	449
R ²	0.511	0.567	0.543	0.605	0.407	0.433
Adjusted R ²	0.507	0.564	0.539	0.601	0.376	0.404

Note: Estimates from regressions predicting each outcome as a function of school and county covariates. All models include school-level fixed effects, a charter indicator, and locale fixed effects. A-F school grade operationalized as a 0-4 GPA variable. Percent possible points is scaled 0-100* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Columns 3–6 show that there are meaningful differences by governance structure. As expected, school and county covariates are much more predictive for TPS than charter school grades, explaining about two-thirds as much of the variation in school grades in charter schools as they do in TPSs. County covariates on their own explain less than half as much of the variation in these measures in charter schools as they do TPSs (7.5% compared with 11-12%, shown Appendix Table A3). Perhaps more surprising, school-level covariates on their own also explain less variation in charters than in TPS schools—about 33–37% in charters compared with 47–52% in TPS (Appendix Table A3). In particular, Table 3 shows that a one standard deviation increase in economically disadvantaged students is associated with a decrease in school grade from 2.8/C+ to about 2.4/C in TPS schools but to just below 2.7/C in charter schools (Columns 3 and 5)—or a decrease on the percent of points index of about 4.6 points in TPSs and 2.2 points in charters.

On the other hand, charter school grades decline more than TPS grades when they serve more students with disabilities and English learners, respectively. For example, a one standard deviation increase in students with disabilities is associated with a decline of just 0.05 grade points and 1.2 percentage points on the percent-of-points index in TPS but 0.26 grade points and 3.1 percentage points in charter schools.

Mechanisms and Alternative Explanations

To the extent that our out-of-school factors are associated with true differences in school quality, the R^2 s in Figure 1 will overstate the role of out-of-school factors in school grades. Our partial mediation models therefore add teacher quality variables, which are school quality measures that are plausibly associated with the out-of-school factors in our models. Table 4 provides the path a estimates from regressions predicting each teacher quality variable as a function of the school and county covariates. Columns 1–3 are from models predicting out-of-field teacher share, 4–6 highly effective teacher share, and 7–9 ineffective teacher share. It is clear that our out-of-school variables are jointly, and in many cases individually, predictive of teacher quality. For example, in the models including both school and county covariates (Columns 3, 6, and 9), a one percentage point increase in Black students is associated with a 0.135 percentage point increase in out-of-field teaching, a 0.31 percentage point decrease in teachers rated highly effective, and a 0.002 percentage point increase in teachers rated unsatisfactory (though this latter estimate is noisy, likely due to the large number of zeroes on the unsatisfactory measure). Because the standard deviation on Black percentage is about 22, that means a one-standard deviation increase in Black students is associated with a 3 percentage point increase in out-of-field teaching and a nearly 7 percentage point decrease in teachers rated highly effective. Other school-level variables follow similarly expected patterns. At the county level, patterns are less straightforward—many of the unexpected results and patterns appear to be driven in part by charters, which tend to be located in urban counties with large Black, Hispanic, and economically disadvantaged populations. Others may stem from county size; for example, more populous counties tend to have greater average educational attainment than less populous counties, and are home to larger schools, which have more teachers and therefore a greater probability of having at least one teacher rated unsatisfactory.

Table 4

Regressions predicting teacher quality variables as a function of school and county covariates (mediation path a)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Out-of-field			Highly effective			Unsatisfactory		
<i>School-level variables</i>									
Economic disadvantage	-0.012 (0.008)		-0.009 (0.009)	-0.361*** (0.023)		-0.124*** (0.025)	-0.004** (0.001)		-0.000 (0.001)
Black	0.131*** (0.010)		0.135*** (0.013)	-0.139*** (0.030)		-0.313*** (0.038)	0.000 (0.002)		0.002 (0.002)
Hispanic	0.049*** (0.012)		0.069*** (0.015)	0.045 (0.036)		0.087* (0.044)	-0.003 (0.002)		-0.001 (0.002)
Asian or Pacific Islander	-0.134** (0.051)		-0.006 (0.052)	0.424** (0.153)		0.130 (0.149)	0.002 (0.008)		-0.005 (0.007)
Indigenous	1.183*** (0.163)		1.007*** (0.163)	0.413 (0.492)		-0.123 (0.462)	-0.032 (0.025)		-0.034 (0.023)
2+ races	0.186* (0.079)		0.307*** (0.084)	0.370 (0.239)		-0.737** (0.237)	0.012 (0.012)		0.007 (0.012)
Students with Disabilities	0.284*** (0.031)		0.216*** (0.032)	0.565*** (0.093)		0.120 (0.090)	0.011* (0.005)		-0.002 (0.004)
English learners	0.121*** (0.020)		0.109*** (0.020)	0.021 (0.059)		-0.273*** (0.056)	0.003 (0.003)		-0.001 (0.003)
<i>County-level variables</i>									
Unemployment Rate		-1.062* (0.502)	-0.941* (0.478)		3.982** (1.443)	4.249** (1.357)		0.188** (0.067)	0.194** (0.068)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Out-of-field			Highly effective			Unsatisfactory		
SNAP Eligibility		-0.409*** (0.083)	-0.214* (0.084)		-1.956*** (0.239)	-1.820*** (0.239)		0.044*** (0.011)	0.042*** (0.012)
Child poverty		0.367** (0.112)	0.377*** (0.107)		1.141*** (0.323)	0.866** (0.302)		0.299*** (0.015)	0.302*** (0.015)
Above Bachelor's		-0.013 (0.056)	0.009 (0.053)		1.552*** (0.162)	1.496*** (0.152)		0.124*** (0.008)	0.123*** (0.008)
Black		0.361*** (0.034)	0.160*** (0.037)		-0.188 (0.097)	0.283** (0.105)		-0.069*** (0.005)	-0.072*** (0.005)
Hispanic		-0.007 (0.029)	-0.075* (0.032)		0.335*** (0.084)	0.316*** (0.089)		-0.050*** (0.004)	-0.049*** (0.004)
Indigenous		0.583 (2.280)	0.146 (2.187)		-4.132 (6.560)	-2.432 (6.206)		0.206 (0.306)	0.214 (0.311)
Asian or Pacific Islander		-1.200*** (0.226)	-1.089*** (0.222)		-5.202*** (0.651)	-5.418*** (0.630)		-0.165*** (0.030)	-0.158*** (0.032)
2+ races		-2.073*** (0.340)	-1.949*** (0.332)		9.677*** (0.979)	10.033*** (0.943)		-0.426*** (0.046)	-0.445*** (0.047)
Constant	-3.577*** (1.042)	18.900*** (4.217)	5.408 (4.140)	81.786*** (3.145)	-22.208 (12.130)	-3.605 (11.748)	0.167 (0.160)	-6.979*** (0.566)	-6.897*** (0.588)
N	2,971	2,971	2,971	2,971	2,971	2,971	2,971	2,971	2,971
R ²	0.213	0.170	0.262	0.230	0.263	0.362	0.022	0.214	0.217
Adjusted R ²	0.209	0.165	0.256	0.227	0.259	0.357	0.017	0.210	0.211

Note: Estimates from regressions predicting each teacher quality variable as a function of school and county covariates. All models include school-level fixed effects, a charter indicator, and locale fixed effects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

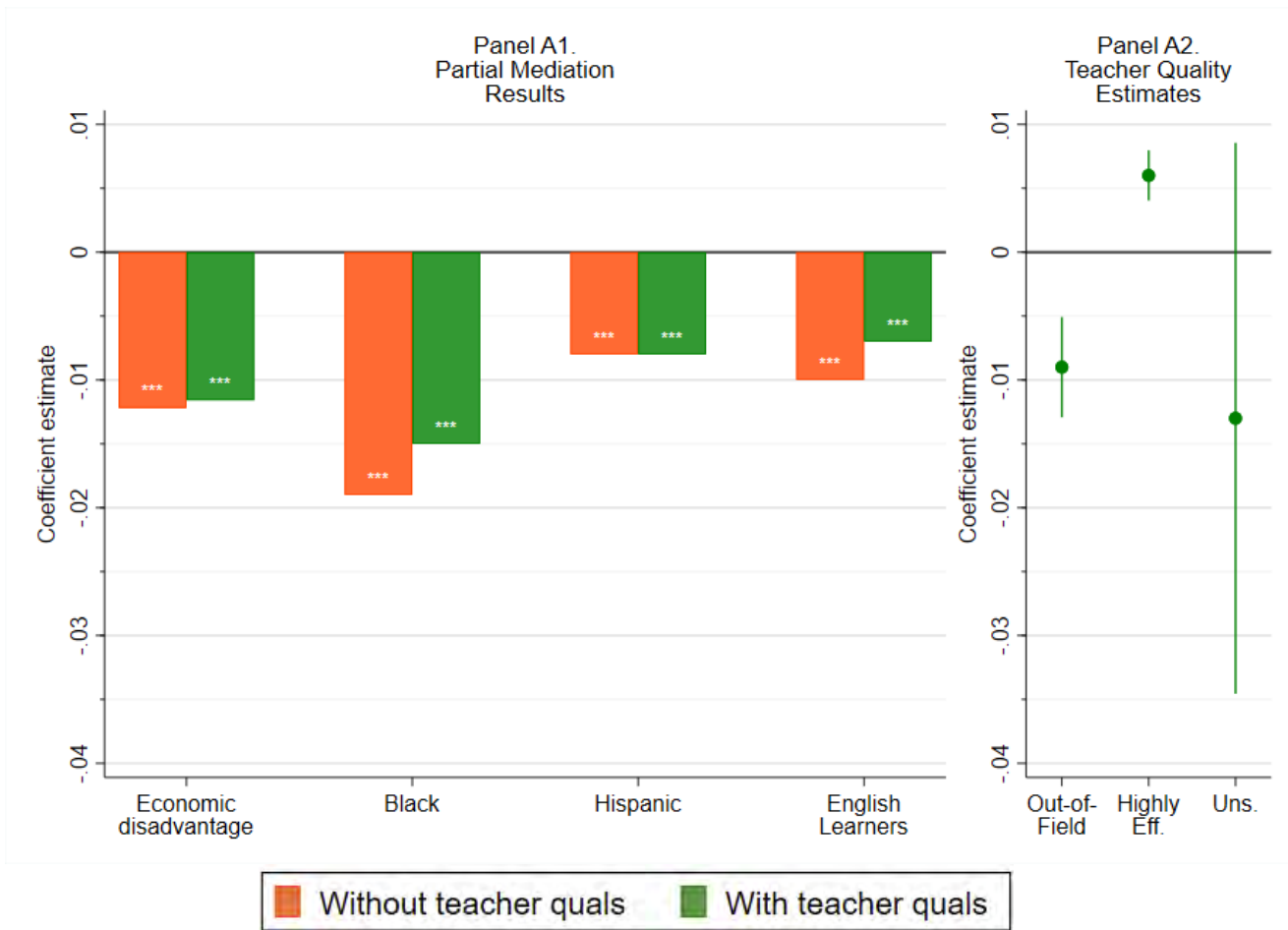
Figure 2 presents a subset of coefficient estimates from the models with and without teacher quality measures, with the A-F grade outcome in Panel A and the percent points outcome in Panel B. Panels A1 and B1 provide the unmediated and mediated estimates, respectively, for four school-level factors that explained a substantial share of the variation in our original models. Panels A2 and B2 display the coefficient estimates (markers) with 95% confidence intervals (spikes) for each of the three teacher quality variables in the mediated model.

There are three takeaways. The first takeaway is that these teacher quality measures are in fact associated with differences in school grades, and in the expected direction, as evidenced by Panels A2 and B2. Having more out-of-field teachers is associated with a decrease in school grade, and having more highly effective teachers is associated with an increase in school grade. Very few teachers are rated unsatisfactory, leading to noisy estimates, but the point estimates there are also negative, showing that having more unsatisfactory teachers is associated with a descriptive decrease in school grade.

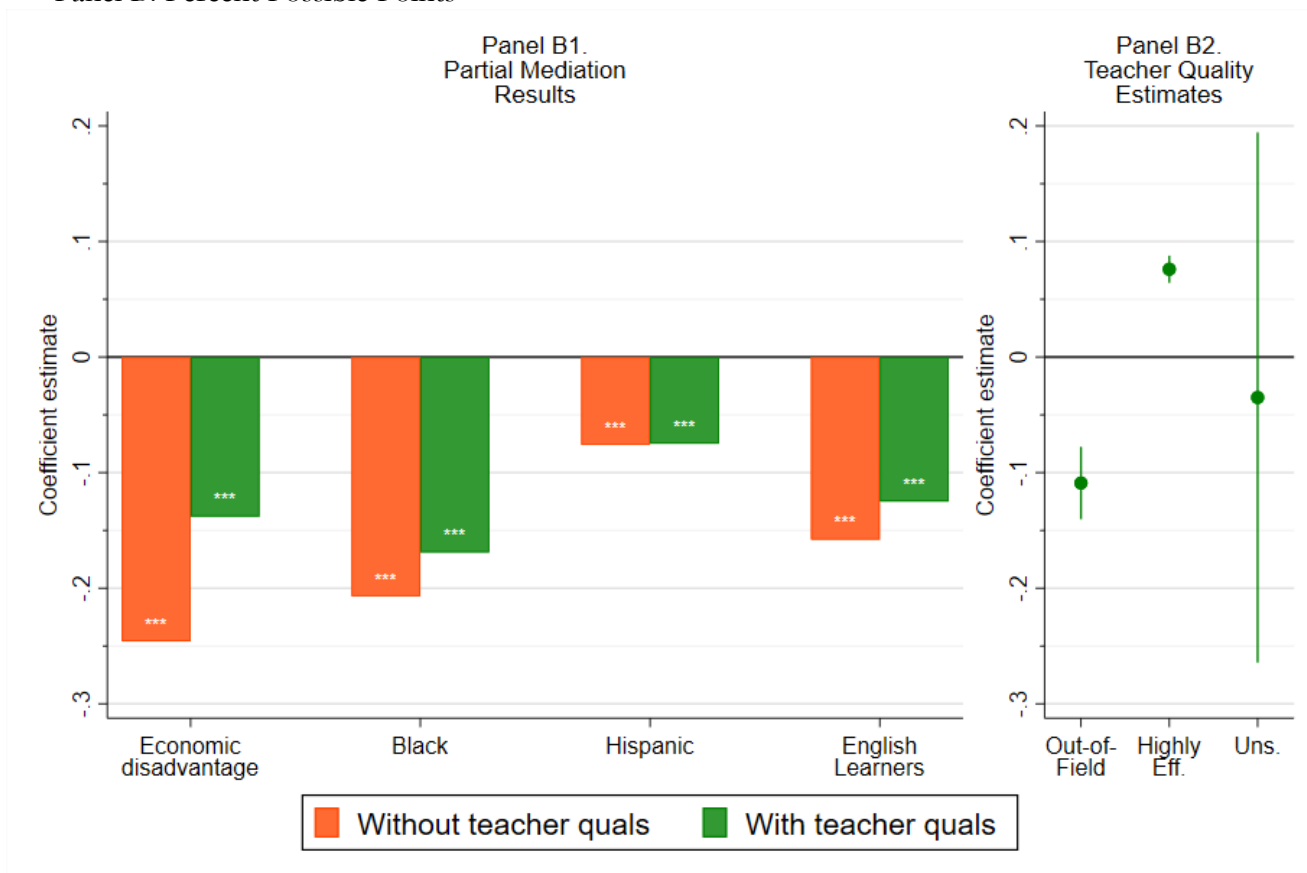
Figure 2

Coefficient Estimates from Models with and without Teacher Quality Variables

Panel A. A-F Grade



Panel B. Percent Possible Points



Note: Panels A1 and B1 provide estimates from unmediated (first/orange bar in bar cluster) and mediated models (second/green bar in bar cluster) predicting A-F school grade in Panel A and 0-100 percent-possible-points index in Panel B. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Panels A2 and B2 provide coefficient estimates and 95% confidence intervals on teacher quality variables in mediated models. Full regression table provided in supplementary material.

The second takeaway is that the teacher quality measures mediate the relationship between some, but not all, socioeconomic/demographic factors and school grade. In the models predicting A-F grade, the coefficient estimate in the mediated model changes on Black and English Learners, but not on economic disadvantage and Hispanic. In the models predicting percent possible points (where there is more variation), the estimate changes on economic disadvantage, Black, and English learners. Together, this suggests that there are true measures of school quality that are associated with some of our out-of-school factors that are not captured in our models. But the third takeaway is the scale is relatively small. For example, inclusion of teacher quality variables attenuated the estimate on Black from -0.019 to -0.015, which means they explain about 20% of the relationship between Black and A-F school grade.⁶ The EL results suggest that teacher quality explains about 30% of the relationship between EL and school grade. The results, in proportional terms, are similar for the percent-possible-points index. Full results from these models are provided in Appendix Table A5.

⁶ This can be calculated as the difference in the coefficient estimates (0.019-0.015=0.004) divided by the unmediated coefficient estimate (0.019).

Finally, Table 5 provides the covariance decomposition results, with the adjusted R^2 from models predicting A-F grade and percent-possible-points index, respectively, as a function of teacher quality covariates only (Column 1/Equation 5), all school and county covariates (Column 2/Equation 4), and the mediated model with all school and county covariates plus the teacher quality covariates (Column 3/Equation 5). In Column 4, we provide the portion of the unmediated adjusted R^2 that covaries with the teacher quality variables, first as an R^2 value and then in parentheses as a percentage of the unmediated adjusted R^2 . This analysis suggests that in total, about 20% of the adjusted R^2 s in the unmediated models can be explained by teacher quality covariates. In other words, about 20% of the variation that our original models attributed to out-of-school factors are confounded with true measures of school quality.

Table 5*Covariance Decomposition*

	(1) Teacher quality variables only (Eq 5)	(2) Unmediated model (Eq 3)	(3) Mediated model (Eq 4)	(4) (C1) – [(C3)-(C2)] (Eq 6)
A-F grade	0.140	0.507	0.539	0.108 (21.3%)
% possible points	0.150	0.562	0.601	0.111 (19.7%)

Note: Cells contain adjusted R^2 s from Equation 5 containing teacher quality variables only (Column 1), Equation 3 containing all school and county covariates (Column 2), and Equation 4 containing all school and county covariates plus the teacher quality mediators (Column 3). The fourth column provides the adjusted R^2 in the unmediated model (i.e., Column 2) that covaries with the teacher quality variables (following Equation 6), followed in parentheses by the percent of the unmediated model adjusted R^2 that is represented (i.e., Column 2 / Column 4).

Discussion

In summary, building from past research on the informational value of school grades (Adams, Forsyth, Ware, & Mwavita, 2016; Adams, Forsyth, Ware, Mwavita, Barnes et al., 2016), we provide the first findings of which we are aware on the extent to which post-pandemic ESSA-era multidimensional school grades can be explained by observable school and county characteristics. This is critically important as states revisit their school grading systems during pandemic recovery, and as federal policymakers consider ESEA reauthorization. We find that all subcomponents of Florida's school grades under ESSA are to some degree explained by school and county covariates—but there is variation by subcomponent. In particular, the proficiency-based subcomponents (especially ELA proficiency) are more thoroughly explained by school and county covariates than the gains-based subcomponents—reiterating a large literature (DeBray et al., 2022; Harbatkin & Wolf, 2023; Heck, 2006; Ho, 2008) that proficiency in particular is a poor measure of school quality. About half of the variation in the multidimensional school grade measure can be explained by our school and county covariates, suggesting that the multidimensional approach to measuring school effectiveness under ESSA appears to capture more meaningful variation than the NCLB era proficiency measures but is still largely confounded by the context the school serves.

Our partial mediation analysis shows that about one-fifth of the variation in the multidimensional index that we attribute to our observed out-of-school factors covaries with true measures of school quality—in particular, teacher qualifications and effectiveness. In other words, after partialing out these teacher quality measures, only about 40% of the school's A-F grade and

45% of the percent-possible-points index is attributable to observed out-of-school factors. Our analyses suggest that teacher quality measures jointly explain from 0-30% of the variation in any given socioeconomic or demographic factor—though most are on the lower end of that range. School demographics such as the share of Black, economically disadvantaged, and English learner students, respectively, are more confounded with teacher quality than others.

This can be interpreted in two ways. The first is that our original models overstate the contribution of out-of-school factors to school grades by about 20% because they misattribute true measures of school quality to demographic and socioeconomic factors. Other unobserved factors that are plausibly measures of school quality, such as school climate, are also associated with socioeconomic and demographic factors we measure (Bryk et al., 2010). The omission of these measures may similarly lead our analyses to overstate the contribution of out-of-school factors to school grades.

However, the second interpretation is that schools serving more Black and economically disadvantaged students and communities, respectively, receive systematically lower grades in large part because they have insufficient access to qualified teachers and other (unobserved) resources. Thus, our findings likely reflect systemic inequalities that lead to the provision of fewer and lower quality educational resources for schools serving marginalized student populations. Penalizing schools for having inadequate access to resources may lead more advantaged families to select away from these schools, resulting in greater segregation, fewer students, and an even greater loss of resources.

In sum, our descriptive findings point to several policy implications. First, consistent with Atchison and colleagues (2023), we find that the subcomponent that measures learning gains for the lowest achieving 25% is least confounded by our school and county characteristics. Research from the NCLB era showed that proficiency-based measures induced a laser focus on students at the margins of proficiency to the detriment of these lower achieving students (e.g., Booher-Jennings, 2005). Together with that work along with literature showing that schools will focus improvement goals on the specific elements that state systems measure (Meyers & VanGronigen, 2019; Mintrop & MacLellan, 2002), this finding suggests that states can induce a focus on the lowest achieving students simply by holding schools accountable for them. In particular, our analysis suggests that performance of the lowest achieving students is substantially less confounded by our set of school and county covariates than are other measures. Thus, including gains of the lowest achievers in an accountability system would have the dual benefit of focusing attention on the learning of these students and providing a school quality measure that appears to be less confounded than others by out-of-school factors. That said, we also caution that a large literature on gaming under NCLB underscores that exclusively holding schools accountable for this subset of students will induce a narrow focus on them to the detriment of other outcomes as well as strategic behavior that is inconsistent with accountability goals. To that end, reporting on growth in other performance quartiles in addition to the bottom quartile would ensure schools remained accountable for the learning of all students while also providing families with a more comprehensive picture of school performance across the distribution.

Second, our analysis suggests that school quality measures appear to be more informative for charters than TPSs. Broadly, this finding points to the possibility that school grades may be more informative in some contexts than others. However, the contributions of charter lottery studies underscore that there are likely unobserved characteristics of families who select into charters that meaningfully contribute to school grades. We also highlight that charter schools in Florida serve fewer English learners and students with disabilities, respectively, than TPSs, and charter school grades appear to be highly sensitive to increasing shares of these students. Still, this finding suggests

that future accountability systems should consider the ways that their design may operate differently for different types of schools.

Third, ESSA does not require that states publicly post school rankings based on its meaningful differentiation index—only that they make ESSA school data publicly available. There is considerable evidence that ratings based on these multidimensional indices contain substantial measurement error and year-to-year volatility (Harbatkin & Wolf, 2023; Hough et al., 2016; Kane & Staiger, 2002; McEachin & Polikoff, 2012). There is also a large and growing literature showing that families are responsive to the way school quality information is presented (Glazerman et al., 2018; Houston & Henig, 2023; Lovenheim & Walsh, 2018; Schneider et al., 2018). Thus, even in the ESSA context—which necessarily includes student achievement subcomponents in the school quality index—states could develop data dashboards that encourage families to draw on more informative measures of school quality, such as gains. They could, for example, choose to report subcomponents separately and privilege the placement of information on more informative subcomponents.

Finally, it is clear from our findings that the multidimensional meaningful differentiation index under ESSA marks an improvement over NCLB's proficiency-based measures because it is less confounded by observable out-of-school factors. However, it is also the case that—at least in Florida—out-of-school factors remain highly predictive of school grades. To that end, more research is needed on how measures can more accurately isolate the effect of schools from the effects of the broader contexts in which they operate.

While this paper is focused largely on implications for policy, it is also the case that district leaders play a crucial role in allocating resources effectively. District leaders could use the bottom 25% measures to plan resource allocation strategies to consider areas of need and promote equity across schools. Finally, there are implications for parents making decisions about where to send their children for school. To the extent that parents are using these school grades to make these decisions, they will overweight out-of-school factors and potentially select into schools that are no more effective than lower rated schools.

There are, of course, several limitations to this analysis as we have highlighted above, and we expect this manuscript to fill only a narrow gap in our understanding of school grades. First, given our focus on post-pandemic, ESSA-era measures, we are limited to a single year of data, which is more sensitive to idiosyncratic year-to-year variation than pooled data would be. Because these data come from the first year of post-pandemic report cards, it is also possible that the association between out-of-school factors and school grades is higher than it would be in other years—though the association may also become stronger if opportunity gaps continue to widen. However, given substantial data limitations in earlier pandemic-affected years, we believe that our findings fill a critical gap in knowledge at a critical period and merit reporting at this stage. That said, we note that our findings align with those of similar studies pre-pandemic, providing some evidence that they are not driven by random chance. For example, a study on Oklahoma's school grades found that differences between letter grades were small and insignificant when student and school characteristics were held constant (Adams, Forysth, Ware, Mwavita, Barnes, & Khojasteh., 2016).

Another limitation arises with the use of Florida's assessment measures. Florida's gains measure is highly specific in its approach to counting students as having “met learning gains” and not as nuanced as, for example, a value-added measure. Like proficiency rates, it also relies on crossing an arbitrary threshold and will therefore necessarily miss movement (gains and losses) away from these thresholds (Ho, 2008). However, many states do dichotomize their growth measures in some way for ESSA (Klein, 2019), so Florida's approach provides a reasonably generalizable “status quo” for analysis. Still, we highlight that our findings would likely look different if they drew on a more nuanced measure of student growth. Finally, while we endeavored to collect a broad spectrum of publicly available variables that are likely to impact school performance, we certainly do not have

a complete census of these measures. Relatedly, though there is significant precedent for using county-level measures to situate school context (e.g., Goldhaber et al., 2022; Harbatkin et al., 2023), the variables we have at the county level capture only a crude community measure because counties tend to be larger than school catchment zones. Ultimately, these two data limitations could lead to an underestimate of the contribution of community characteristics to school grades.

Together, our findings as well as the limitations of our analyses point to several avenues for future research. First, future research should draw on multiple years of data to unpack the extent to which these school and community factors appear to contribute to school grades—and how that may change as we move further from the pandemic’s onset. Other research has shown that the pandemic wrought outsized negative effects on the lowest performing schools and the communities they serve (Harbatkin, Strunk, et al., 2023); it would be useful to know whether the contribution of school and community factors attenuates, grows stronger, or remains stable in future years. Second, Florida’s standing as a pioneer in school accountability policy makes it a useful context through which to ask these research questions, but additional research in other state contexts could help to uncover the extent to which different state approaches to measuring school quality may be more or less effective at capturing the teaching and learning that occurs within schools rather than the community contexts in which schools operate. Third, our math versus ELA findings buttresses a large literature showing that intervention effects tend to be larger in math than ELA by demonstrating empirically that ELA performance is driven more by out-of-school factors than math performance. Future research could consider the mechanisms through which these out-of-school factors appear to influence ELA versus math performance.

References

- Adams, C. M., Forsyth, P. B., Ware, J., & Mwavita, M. (2016). The informational significance of A–F school accountability grades. *Teachers College Record*, *118*(7), 1–31. <https://doi.org/10.1177/016146811611800707>
- Adams, C. M., Forsyth, P. B., Ware, J. K., Mwavita, M., Barnes, L. L., & Khojasteh, J. (2016). An empirical test of Oklahoma’s A-F grades. *Education Policy Analysis Archives*, *24*, 4–4. <https://doi.org/10.14507/epaa.v24.2127>
- Atchison, D., Blair, D., Hyland, K., Ozek, U., & Floch, K. L. (2023). *Identification of comprehensive support and improvement schools in Florida*. https://www.air.org/sites/default/files/2024-04/23-23295-Florida-Measures-Report_Dec2023_fmt_ed120723-Clean.pdf
- Balfanz, R., Legters, N., West, T. C., & Weber, L. M. (2007). Are NCLB’s measures, incentives, and improvement strategies the right ones for the nation’s low-performing high schools? *American Educational Research Journal*, *44*(3), 559–593. <https://doi.org/10.3102/0002831207306768>
- Ballou, D., & Springer, M. G. (2016). Has NCLB encouraged educational triage? Accountability and the distribution of achievement gains. *Education Finance and Policy*, *12*(1), 77–106. https://doi.org/10.1162/EDFP_a_00189
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bayer, P., Ferreira, F., & McMillan, R. (2004). *Tiebout sorting, social multipliers and the demand for school quality* (Working Paper 10871). National Bureau of Economic Research. <https://doi.org/10.3386/w10871>
- Bernal, P., Mittag, N., & Qureshi, J. A. (2016). Estimating effects of school quality using multiple proxies. *Labour Economics*, *39*, 1–10. <https://doi.org/10.1016/j.labeco.2016.01.005>

- Beuermann, D. W., Jackson, C. K., Navarro-Sola, L., & Pardo, F. (2023). What is a good school, and can parents tell? Evidence on the multidimensionality of school output. *The Review of Economic Studies*, 90(1), 65–101. <https://doi.org/10.1093/restud/rdac025>
- Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education*. *The Quarterly Journal of Economics*, 114(2), 577–599. <https://doi.org/10.1162/003355399556070>
- Bonilla, S., & Dee, T. (2020). The effects of school reform under NCLB waivers: Evidence from focus schools in Kentucky. *Education Finance and Policy*, 15(1), 75–103. https://doi.org/10.1162/edfp_a_00275
- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas accountability system. *American Educational Research Journal*, 42(2), 231–268. <https://doi.org/10.3102/00028312042002231>
- Bryk, A., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. The University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/O/bo8212979.html>
- Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3), 535–576. <https://doi.org/10.1086/669340>
- Burns, J., Harbatkin, E., Strunk, K. O., Torres, C., Mcilwain, A., & Frost Waldron, S. (2023). The efficacy and implementation of Michigan’s Partnership Model of school and district turnaround: Mixed-methods evidence from the first 2 years of reform implementation. *Educational Evaluation and Policy Analysis*, 45(4), 622–654. <https://doi.org/10.3102/01623737221141415>
- Candelaria, C. A., & Shores, K. A. (2019). Court-ordered finance reforms in the adequacy era: Heterogeneous causal effects and sensitivity. *Education Finance and Policy*, 14(1), 31–60. https://doi.org/10.1162/edfp_a_00236
- Carlson, D., & Lavertu, S. (2018). School improvement grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis*, 0162373718760218. <https://doi.org/10.3102/0162373718760218>
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star *. *The Quarterly Journal of Economics*, 126(4), 1593–1660. <https://doi.org/10.1093/qje/qjr041>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, 104(9), 2633–2679. <https://doi.org/10.1257/aer.104.9.2633>
- Cyrus, E., Clarke, R., Hadley, D., Bursac, Z., Trepka, M. J., Dévieux, J. G., Bagci, U., Furr-Holden, D., Coudray, M., Mariano, Y., Kiplagat, S., Noel, I., Ravelo, G., Paley, M., & Wagner, E. F. (2020). The impact of COVID-19 on African American communities in the United States. *Health Equity*, 4(1), 476–483. <https://doi.org/10.1089/heq.2020.0030>
- Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of ‘No Child Left Behind.’ *Race Ethnicity and Education*, 10(3), 245–260. <https://doi.org/10.1080/13613320701503207>
- Darling-Hammond, L. (2018). From “separate but equal” to “No Child Left Behind”: The collision of new standards and old inequalities. In *Thinking about schools* (pp. 419–437). Routledge.
- Darling-Hammond, L., & Snyder, J. (2015). Accountability for resources and outcomes: An introduction. *Education Policy Analysis Archives*, 23, 20–20. <https://doi.org/10.14507/epaa.v23.2024>
- Davis, T., Bhatt, R., & Schwarz, K. (2015). School segregation in the era of accountability. *Social Currents*, 2(3), 239–259. <https://doi.org/10.1177/2329496515589852>

- DeBray, E., Finnigan, K. S., George, J., & Scott, J. (2022). *A civil rights framework for the Reauthorization of ESEA*. National Education Policy Center.
- Dee, T., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418–446. <https://doi.org/10.1002/pam.20586>
- Dee, T., Jacob, B., & Schwartz, N. L. (2013). The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis*, 35(2), 252–279. <https://doi.org/10.3102/0162373712467080>
- Denice, P., & Gross, B. (2016). Choice, preferences, and constraints: Evidence from public school applications in Denver. *Sociology of Education*, 89(4), 300–320. <https://doi.org/10.1177/0038040716664395>
- Dijkstra, A., Daas, R., De la Motte, P., & Ehren, M. (2017). Inspecting school social quality: Assessing and improving school effectiveness in the social domain. *Journal of Social Science Education*, 16(4), Article 4.
- Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, 32(4), 692–717. <https://doi.org/10.1002/pam.21715>
- Education Commission of the States. (2021, December). *50-state comparison: States' school accountability systems*. <https://www.ecs.org/50-state-comparison-states-school-accountability-systems/>
- Engel, M., Jacob, B. A., & Curran, F. C. (2014). New evidence on teacher labor supply. *American Educational Research Journal*, 51(1), 36–72. <https://doi.org/10.3102/0002831213503031>
- Eren, O., Figlio, D. N., Mocan, N. H., & Ozturk, O. (2023). *School accountability, long-run criminal activity, and self-sufficiency* (Working Paper 31556). National Bureau of Economic Research. <https://doi.org/10.3386/w31556>
- Figlio, D., & Loeb, S. (2011). School accountability. *Handbook of the Economics of Education*, 3, 383–421. <https://doi.org/10.1016/B978-0-444-53429-3.00008-9>
- Figlio, D. N., & Lucas, M. E. (2004). What's in a grade? School report cards and the housing market. *American Economic Review*, 94(3), 591–604. <https://doi.org/10.1257/0002828041464489>
- Finch, W. H., & Hernández Finch, M. E. (2020). Poverty and COVID-19: Rates of incidence and deaths in the United States during the first 10 weeks of the pandemic. *Frontiers in Sociology*, 5. <https://doi.org/10.3389/fsoc.2020.00047>
- Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal*, 44(3), 594–630. <https://doi.org/10.3102/0002831207306767>
- Florida Department of Education. (2021, July). *2020-21 guide to calculating school grades and district grades*. <https://www.fldoe.org/core/fileparse.php/18534/urlt/SchoolGradesCalcGuide21.pdf>
- Frankenberg, E. (2018). Preferences, proximity, and controlled choice: Examining families' school choices and enrollment decisions in Louisville, Kentucky. *Peabody Journal of Education*, 93(4), 378–394. <https://doi.org/10.1080/0161956X.2018.1488392>
- Fusarelli, L. D. (2004). The potential impact of the No Child Left Behind Act on equity and diversity in American education. *Educational Policy*, 18(1), 71–94. <https://doi.org/10.1177/0895904803260025>
- Gamoran, A. (2008). *Standards-based reform and the poverty gap: Lessons for "No Child Left Behind"*. Rowman & Littlefield.
- Gamoran, A. (2015). *The future of educational inequality in the United States: What went wrong, and how can we fix it?* William T. Grant Foundation. <http://wtgrantfoundation.org/resource/the-future-of-educational-inequality-what-went-wrong-and-how-can-we-fix-it>
- Garcia, D. R. (2008). The impact of school choice on racial segregation in charter schools. *Educational Policy*, 22(6), 805–829. <https://doi.org/10.1177/0895904807310043>

- Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy, 11*(2), 125–149. https://doi.org/10.1162/EDFP_a_00180
- Glazerman, S., Nichols-Barrer, I., Valant, J., Chandler, J., Burnett, A., et al. (2018). *Nudging parents to choose better schools: The importance of school choice architecture*. Mathematica Policy Research.
- Goldhaber, D., Imberman, S. A., Strunk, K. O., Hopkins, B. G., Brown, N., Harbatkin, E., & Kilbride, T. (2022). To what extent does in-person schooling contribute to the spread of COVID-19? Evidence from Michigan and Washington. *Journal of Policy Analysis and Management, 41*(1), 318–349. <https://doi.org/10.1002/pam.22354>
- Gregg, J. J., & Lavertu, S. (2023). Test-based accountability and educational equity: Breaking through local district politics? *Economics of Education Review, 97*, 102485. <https://doi.org/10.1016/j.econedurev.2023.102485>
- Hamlin, D. (2021). Can a positive school climate promote student attendance? Evidence From New York City. *American Educational Research Journal, 58*(2), 315–342. <https://doi.org/10.3102/0002831220924037>
- Harbatkin, E., Nguyen, T. D., Strunk, K. O., Burns, J., & Moran, A. (2023). *Should I stay or should I go (later)? Teacher intentions and turnover in low-performing schools and districts before and during the COVID-19 pandemic*. Annenberg Institute at Brown University. <https://doi.org/10.26300/d7dh-kq82>
- Harbatkin, E., Pham, L. D., Redding, C., & Moran, A. J. (2024). What are the side effects of school turnaround? A systematic review. *Review of Research in Education, 47*(1). <https://doi.org/10.3102/0091732X241248151>
- Harbatkin, E., Strunk, K. O., & McIlwain, A. (2023). School turnaround in a pandemic: An examination of the outsized implications of COVID-19 on low-performing turnaround schools, districts, and their communities. *Economics of Education Review, 97*, 102484. <https://doi.org/10.1016/j.econedurev.2023.102484>
- Harbatkin, E., & Wolf, B. (2023). *State accountability decisions under the Every Student Succeeds Act and the validity, stability, and equity of school ratings*. Annenberg Institute at Brown University. <https://doi.org/10.26300/xt8e-0w18>
- Hasan, S., & Kumar, A. (2019). Digitization and divergence: Online school ratings and segregation in America. *SSRN, 3265316*. <https://doi.org/10.2139/ssrn.3265316>
- Hashim, S. A., Kelley-Kemple, T., & Laski, M. E. (2023). *An improved method for estimating school-level characteristics from census data*. Annenberg Institute at Brown University. <https://www.edworkingpapers.com/ai23-804>
- Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics, 123*(4), 1373–1414. <https://doi.org/10.1162/qjec.2008.123.4.1373>
- Heck, R. H. (2006). Assessing school achievement progress: Comparing alternative approaches. *Educational Administration Quarterly, 42*(5), 667–699. <https://doi.org/10.1177/0013161X06293718>
- Ho, A. D. (2008). The problem with “proficiency”: Limitations of statistics and policy under No Child Left Behind. *Educational Researcher, 37*(6), 351–360. <https://doi.org/10.3102/0013189X08323842>
- Hough, H., Penner, E., & Witte, J. (2016). *Identity crisis: Multiple measures and the identification of schools under ESSA*. Policy Memo 16-3. Policy Analysis for California Education, PACE. <https://eric.ed.gov/?id=ED574851>
- Houston, D. M., & Henig, J. R. (2023). The “good” schools: Academic performance data, school choice, and segregation. *AERA Open, 9*, 23328584231177666. <https://doi.org/10.1177/23328584231177666>

- Ingersoll, R. (2004). *Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers?* Renewing Our Schools, Securing Our Future - A National Task Force on Public Education; Joint Initiative of the Center for American Progress and the Institute for America's Future. https://repository.upenn.edu/gse_pubs/493
- Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, 38(3), 499–534. <https://doi.org/10.3102/00028312038003499>
- Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics*, 27(2), 213–256. <https://doi.org/10.1086/599334>
- Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (Working Paper 18624). National Bureau of Economic Research. <https://doi.org/10.3386/w18624>
- Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, 131(1), 157–218. <https://doi.org/10.1093/qje/qjv036>
- Jackson, C. K., & Mackevicius, C. L. (2024). What impacts can we expect from school spending policy? Evidence from evaluations in the United States. *American Economic Journal: Applied Economics*, 16(1), 412–446. <https://doi.org/10.1257/app.20220279>
- Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, 16(4), 91–114. <https://doi.org/10.1257/089533002320950993>
- Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher*, 34(8), 3–13. <https://doi.org/10.3102/0013189X034008003>
- Kitzmiller, E. M. (2020). “We are the forgotten of the forgotten”: The effects of charter school reform on public school teachers. *Harvard Educational Review*, 90(3), 371–396. <https://doi.org/10.17763/1943-5045-90.3.371>
- Klein, A. (2019, January 23). How are states measuring student growth under ESSA? *Education Week*. <https://www.edweek.org/education/how-are-states-measuring-student-growth-under-essa/2019/01>
- Kotok, S., Frankenberg, E., Schafft, K. A., Mann, B. A., & Fuller, E. J. (2017). School choice, racial segregation, and poverty concentration: Evidence From Pennsylvania charter school transfers. *Educational Policy*, 31(4), 415–447. <https://doi.org/10.1177/0895904815604112>
- Krieg, J. M., & Storer, P. (2006). How much do students matter? Applying the Oaxaca decomposition to explain determinants of adequate yearly progress. *Contemporary Economic Policy*, 24(4), 563–581. <https://doi.org/10.1093/cep/byl003>
- Ladd, H. F. (2017). No Child Left Behind: A deeply flawed federal policy. *Journal of Policy Analysis and Management*, 36(2), 461–469. <https://doi.org/10.1002/pam.21978>
- Le Floch, K., Atchison, D., Ozek, U., Hyland, K., Blair, D., & Hurlburt, S. (2023). *Multiple measure accountability under ESSA: Early findings from three states*. https://www.air.org/sites/default/files/2024-04/23-20801-CSI-NCER-ESSA-measures-brief-FMT-ed_rev.pdf
- Lee, J., & Lubienski, C. (2017). The impact of school closures on equity of access in Chicago. *Education and Urban Society*, 49(1), 53–80. <https://doi.org/10.1177/0013124516630601>
- Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990–2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, 34(2), 209–231. <https://doi.org/10.3102/0162373711431604>

- Lenzi, M., Vieno, A., Sharkey, J., Mayworm, A., Scacchi, L., Pastore, M., & Santinello, M. (2014). How school can teach civic engagement besides civic education: The role of democratic school climate. *American Journal of Community Psychology*, *54*(3), 251–261. <https://doi.org/10.1007/s10464-014-9669-8>
- Lin, A. (2015). Citizenship education in American schools and its role in developing civic engagement: A review of the research. *Educational Review*, *67*(1), 35–63. <https://doi.org/10.1080/00131911.2013.813440>
- Lipman, P. (2017). The landscape of education “reform” in Chicago: Neoliberalism meets a grassroots movement. *Education Policy Analysis Archives*, *25*, 54–54. <https://doi.org/10.14507/epaa.25.2660>
- Lovenheim, M. F., & Walsh, P. (2018). Does choice increase information? Evidence from online school search behavior. *Economics of Education Review*, *62*, 91–103. <https://doi.org/10.1016/j.econedurev.2017.11.002>
- Mansfield, J., & Slichter, D. (2021). *The long-run effects of consequential school accountability*. IZA Institute of Labor Economics. <https://doi.org/10.2139/ssrn.3879350>
- McEachin, A., & Polikoff, M. S. (2012). We are the 5%: Which schools would be held accountable under a proposed revision of the Elementary and Secondary Education Act? *Educational Researcher*, *41*(7), 243–251. <https://doi.org/10.3102/0013189X12453494>
- Meyers, C. V., & VanGronigen, B. A. (2019). A lack of authentic school improvement plan development: Evidence of principal satisficing behavior. *Journal of Educational Administration*, *57*(3), 261–278. <https://doi.org/10.1108/JEA-09-2018-0154>
- Mintrop, H., & MacLellan, A. M. (2002). School improvement plans in elementary and middle schools on probation. *The Elementary School Journal*, *102*(4), 275–300. <https://doi.org/10.1086/499704>
- Owens, A., Reardon, S. F., & Jencks, C. (2016). Income segregation between schools and school districts. *American Educational Research Journal*, *53*(4), 1159–1197. <https://doi.org/10.3102/0002831216652722>
- Owens, A., & Sunderman, G. L. (2006). *School accountability under NCLB: Aid or obstacle for measuring racial equity?* Civil Rights Project. <https://escholarship.org/content/qt9sv829ng/qt9sv829ng.pdf>
- Pearman, F. A., & Marie Greene, D. (2022). School closures and the gentrification of the Black metropolis. *Sociology of Education*, *95*(3), 233–253. <https://doi.org/10.1177/00380407221095205>
- Pivovarova, M., & Powers, J. (2024, March 16). *The practical and policy relevance of growth-focused school grades*. Association for Education Finance and Policy. <https://virtual.oxfordabstracts.com/#/event/4542/submission/539>
- Reardon, S. F. (2019). *Affluent schools are not always the best schools*. The Educational Opportunity Project at Stanford University. <https://edopportunity.org/discoveries/affluent-schools-are-not-always-best/>
- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, *92*(5), 1394–1415. <https://doi.org/10.1016/j.jpubeco.2007.05.003>
- Redding, C., & Nguyen, T. D. (2020). Recent trends in the characteristics of new teachers, the schools in which they teach, and their turnover rates. *Teachers College Record*, *122*(7), 1–36. <https://doi.org/10.1177/016146812012200711>
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>

- Robertson, J. S., Smith, R. W., & Rinka, J. (2016). How did successful high schools improve their graduation rates? *Journal of At-Risk Issues*, 19(1), 10–18.
- Rothstein, J. M. (2006). Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *The American Economic Review*, 96(4), 1333–1350. <https://doi.org/10.1257/aer.96.4.1333>
- Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2013). Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure. *American Economic Journal: Economic Policy*, 5(2), 251–281. <https://doi.org/10.1257/pol.5.2.251>
- Schneider, J., Jacobsen, R., White, R., & Gehlbach, H. (2017). Building a better measure of school quality. *Phi Delta Kappan*, 98(7), 43–48. <https://doi.org/10.1177/0031721717702631>
- Schneider, J., Jacobsen, R., White, R. S., & Gehlbach, H. (2018). The (mis)measure of schools: How data affect stakeholder knowledge and perceptions of quality. *Teachers College Record*, 120(5), 1–40. <https://doi.org/10.1177/016146811812000507>
- Strunk, K. O., Marsh, J. A., Hashim, A. K., & Bush-Mecenas, S. (2016). Innovation and a return to the status quo: A mixed-methods study of school reconstitution. *Educational Evaluation and Policy Analysis*. <https://doi.org/10.3102/0162373716642517>
- Sun, M., Kennedy, A. I., & Loeb, S. (2021). The longitudinal effects of school improvement grants. *Educational Evaluation and Policy Analysis*, 43(4), 647–667. <https://doi.org/10.3102/01623737211012440>
- Sun, M., Penner, E. K., & Loeb, S. (2017). Resource- and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal*, 54(4), 607–643. <https://doi.org/10.3102/0002831217695790>
- Winters, M. A., & Cowen, J. M. (2012). Grading New York: Accountability and student proficiency in America's largest school district. *Educational Evaluation and Policy Analysis*, 34(3), 313–327. <https://doi.org/10.3102/0162373712440039>

About the Authors

Nandrea Burrell

Florida State University

burrell@psy.fsu.edu

<https://orcid.org/0000-0003-2824-4755>

Nandrea Burrell is a developmental psychology PhD candidate at Florida State University. Her research examines contextual factors that promote or hinder children's educational attainment.

Erica Harbatkin

Florida State University

eharbatkin@fsu.edu

<https://orcid.org/0000-0001-8304-2502>

Erica Harbatkin is an assistant professor of educational policy and evaluation at Florida State University. Her research focuses on school accountability and improvement, teacher policy, and educational equity with an emphasis on the role of public policy in shaping student outcomes.

education policy analysis archives

Volume 32 Number 32

July 2, 2024

ISSN 1068-2341



Readers are free to copy, display, distribute, and adapt this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, the changes are identified, and the same license applies to the derivative work. More details of this Creative Commons license are available at <https://creativecommons.org/licenses/by-sa/4.0/>. **EPAA** is published by the Mary Lou Fulton Teachers College at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A1 (Brazil), SCImago Journal Rank, SCOPUS, Socolar (China).

About the Editorial Team: <https://epaa.asu.edu/ojs/index.php/epaa/about/editorialTeam>

Please send errata notes to Jeanne M. Powers at jeanne.powers@asu.edu
