The Lasting Impacts of Middle School Principals

by Eric A. Hanushek, Andrew Morgan, Steven G. Rivkin, Jeffrey C. Schiman, Ayman Shakeel, and Lauren Sartain [1]

October 2024

ABSTRACT

Using rich Texas administrative data, we estimate the impact of middle school principals on post-secondary schooling, employment, and criminal justice outcomes. The results highlight the importance of school leadership, though striking differences emerge in the relative importance of different skill dimensions to different outcomes. The estimates reveal large and highly significant effects of principal value-added to cognitive skills on the productive activities of schooling and work but much weaker effects of value-added to noncognitive skills on these outcomes. In contrast, there is little or no evidence that middle school principals affect the probability a male is arrested and has a guilty disposition by raising cognitive skills but strong evidence that they affect these outcomes through their impacts on noncognitive skills, especially those related to the probability of an out-of-school suspension. Heterogeneity analysis reveals that the principal effect on the probability a male is arrested is strongest for males with the highest predicted risk of arrest based on information prior to middle school entry, while principal effects on the probability of attending and persisting in college span the predicted risk distributions outside of the top decile. Finally, the principal effects on the probability of engagement in the criminal justice system are much larger for Black than for nonblack males.

[1] Hanushek: Stanford University, National Bureau of Economic Research, and University of Texas at Dallas; Morgan: University of Illinois Chicago: Rivkin: University of Illinois at Chicago, National Bureau of Economic Research and University of Texas at Dallas; Schiman, Georgia Southern University; Shakeel, University of Illinois Chicago; and Sartain, University of North Carolina; This research has been supported by the Laura and John Arnold Foundation, the Smith Richardson Foundation, and CALDER. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, or the State of Texas.

I. Introduction

Effective leadership is often assumed to be a key element of successful schools, but conceptual and data limitations have complicated the identification of the principal's contribution to school quality and hampered the assessment of this assumption. A growing body of research tackles these impediments, and we build on this work in our study of middle school principal effects on a range of post-secondary outcomes including college attendance and persistence, engagement with the productive activities of schooling and work, and interactions with the criminal justice system. We find that middle school principals exert strong impacts on these later-life outcomes through their effects on the development of both cognitive and noncognitive skills.

We use a two-stage approach to identify principal impacts on longer-term outcomes that builds on studies of teacher and school effects by Chetty, Friedman, and Rockoff (2014a), Jackson (2018), and Jackson et al. (2020). We first estimate principal value-added to the development of cognitive and noncognitive skills, where absences and out-of-school suspensions serve as the primary proxies for two potentially important dimensions of noncognitive skills and average math and reading test scores serve as the proxy for cognitive skills. We then estimate the relationships between principal value-added and longer-term outcomes. Our empirical framework mitigates the potentially confounding influences introduced by nonrandom school sorting of students and educators, and it also directly addresses confounding static and dynamic school factors that potentially bias estimates of principal effectiveness.

Isolating the contribution of principals presents a formidable analytical challenge, similar to that for the estimation of teacher effects. Both must address the endogeneity of family selection of neighborhoods and schools and the fact that skill proxies may be subject to manipulation in response to school accountability pressures. A major difference between the two is the presence in each school of only a single principal as opposed to multiple teachers at any point in time. Although this removes concerns related to the purposeful placement of students into classrooms, it amplifies the difficulty of

accounting for time-varying neighborhood and school factors out of the principal's control.

The availability of rich administrative data enables the use of panel data methods that account for student heterogeneity and variation over time in labor market and criminal justice conditions, district finances, leadership, curriculum and disciplinary policies, and neighborhood demographics. In addition, the size and diversity of Texas and length of the panel data sets reduce sample attrition and support an analysis of heterogeneity by race, an investigation that is particularly important given racial differences in involvement with the criminal justice system.

We find that principals affect post-secondary academic, labor market, and crime outcomes through their influences on the development of both cognitive and noncognitive skills. Principal value-added to cognitive skills is strongly related to the probability of engaging in the productive activities of college attendance, college persistence, and work. The primary impact in the noncognitive domain comes through engagement with the criminal justice system. Value-added to both absences and the probability of receiving an out-of-school suspension significantly influence the probabilities that a male is ever arrested and has a guilty disposition.[1]

The proxies researchers have used to measure noncognitive skills vary across studies, and there remains considerable debate over information captured by them. Due to data availability, we use absences and suspensions as proxies for two dimensions of noncognitive skills – and this may contribute to the finding that noncognitive skill effects are most important in the case of criminal justice outcomes. The positive correlations between principal effects on achievement and those on both absences and suspensions suggest that principals who are more effective at raising cognitive skills may also be more effective at raising noncognitive skills more strongly related to engagement in productive activities.

[1] Rates of arrest are far lower for women than for men, and therefore the criminal justice component focuses on males.

Heterogeneity analysis reveals that principal effect on the probability a male is arrested is strongest for males with the highest predicted risk of arrest based on information prior to middle school entry, while principal effects on the probability of attending and persisting in college span the predicted risk distributions outside of the top decile. Examination of heterogeneity by race finds race differences in the patterns of principal effects on the probability a male is arrested and the probability of attending and persisting in college: there is a much stronger relationship between the probability of arrest and suspension VA for Black students but stronger relationships between probability of arrest and absence VA for Hispanic and especially for White students. Effects of principal VA to test score on college attendance and persistence are strongest for Hispanic students, somewhat less strong for Black students, and small and insignificant for the remaining students.

We provide detailed evidence about the sensitivity of the estimates to the methods of controlling for confounding influences, to the treatment of outmigration from Texas, to random cohort shocks, and to whether skill proxy measures are subject to strategic manipulation in response to accountability pressures. The estimation of multiple dimensions of principal effects using methods that account for student heterogeneity and static and time-varying school influences paints a comprehensive picture of the importance of school leadership.

The next section reviews the research upon which we build our empirical models. Section 3 describes the Texas administrative data, and Section 4 develops the specifications used to identify principal value-added to skills and the relationships between outcomes and the value-added estimates. Section 5 presents a series of estimates of the effects of value-added to skills on the post-secondary schooling, work, and criminal justice outcomes. An analysis of heterogeneous effects follows the presentation of results for all students combined. The final section summarizes the analysis and considers implications for the measurement of principal productivity and policy.

II. Prior Research

The expansion of school accountability has led to increased emphasis on the quality of school leadership. A growing body of research investigates the principal's contributions to variation in student outcomes, similar to studies on schools and teachers including the aforementioned work by Chetty, Friedman, and Rockoff (2014a), Jackson (2018) and Jackson et al. (2020).[2] Much of the research on principal productivity focuses on achievement, and there is ongoing debate on the identification of principal effects. Laing et al. (2016) introduced concerns about potentially confounding time-varying school factors, and these were also highlighted in Bartanen et al. (2024). However, both Laing et al. (2016) and Branch et al. (2020) find strong evidence of significant variation in principal productivity once specifications account for both time-varying school influences and disruptions around principal transitions.[3]

Related studies of the impacts of teachers and schools on both immediate and longer-term outcomes inform the structure of our analysis.[4] The limitations of a singular focus on achievement are clear from evidence that teacher effects on both cognitive and noncognitive skills contribute to longer-term academic, social, and labor-market outcomes.[5] Because of the complexity of educational dynamics and the possibility of confounding intervening factors, we adopt a version of the two-step estimation approaches pioneered for teacher and school value-added by Chetty, Friedman, and Rockoff (2014a), Jackson (2018), and Jackson et al. (2020). We estimate principal value-

[2] Studies include Jacob and Lefgren (2008), Clark, Martorell, and Rockoff (2009), Branch, Hanushek, and Rivkin (2012), Chiang, Lipscomb, and Gill (2016), Hochbein and Cunningham (2013), Dhuey and Smith (2014), and Grissom, Kalogrides, and Loeb (2015).

[3] Research also raises other specification concerns. For example, Grissom, Kalogrides, and Loeb (2015) emphasize the sensitivity of estimates to the empirical framework. Branch et al. (2020) address this concern as well.

[4] There are now a number of reviews of research into value-added of teachers; see, for example, Hanushek and Rivkin (2012), Koedel, Mihaly, and Rockoff (2015), Bacher-Hicks and Koedel (2023).

[5] The growing body of research on teacher and school effects on long-term outcomes includes the studies by Card and Krueger (1992a, (1992b) on school resource effects, Elango, García, Heckman, and Hojman (2016) on the effects of early schooling interventions, Chetty, Friedman, and Rockoff (2014a, (2014b) on the effects of teacher quality on earnings, Jackson (2018) on the effects of teachers on high school completion and other longer-term outcomes, Jackson et al. (2020) on the development of socio-emotional skills and their effects on crime and educational attainment, and Deming, Cohodes, Jennings, and Jencks (2016) on accountability effects. Petek and Pope (2023) show evidence on how multidimensional teacher value-added affects subsequent student outcomes.

added to cognitive and non-cognitive skills using proximate schooling data and then relate these estimates to the subsequent post-secondary outcomes.

Research by Carrell and West (2010), Gilraine and Pope (2021), and Dinerstein and Opper (2022) highlights the importance of focusing on longer-term skill measures when estimating educator value-added, particularly when the short-term outcomes are high-stakes for the educator. Gilraine and Pope (2021) provide evidence that measuring teacher effectiveness using longer-term, low-stakes test scores elevates the importance of cognitive skills relative to noncognitive skills and is more closely related to future education, economic, and social outcomes than effectiveness at raising short-term measures. Dinerstein and Opper (2022), in a partial-observability model, demonstrate that educators will respond to policies that elevate the importance of end-of-year test scores as opposed to other unmeasured outcomes. These studies lead us to use skills measured in 9th grade following matriculation to high school in the estimation of middle-school principal effectiveness. Doing so lessens the potential influence of strategic responses to accountability pressures and, along with high school-by-cohort fixed effects, mitigates complications introduced by differing standards and operating procedures across schools for the issuance of disciplinary infractions and out-of-school suspensions.

The importance of considering criminal justice outcomes along with school attainment and labor market performance becomes clear from studies of arrests and incarceration. Lochner and Moretti (2004) establish the negative effect of school attainment on the probabilities of arrest and incarceration, and other research points to the important role of acquired noncognitive skills. Rose, Schellenberg, and Shem-Tov (2022) find evidence that a higher teacher value-added to noncognitive skills (but not to cognitive skills) reduces the probability of an arrest. In two studies of Charlotte-Mecklenburg middle schools, Deming (2011) and Bacher-Hicks, Billings, and Deming (2019) use exogenous variation produced by school choice lotteries and unexpected changes in school attendance zones, respectively, to identify the effects of middle school quality and suspension rates on the probability of engagement with the criminal justice system as an adult. Both find a positive effect of suspensions on the probability of engagement with the criminal justice system, though they do not attempt to disentangle the channels of cognitive and noncognitive skill development, to identify the effects of

principals, or to estimate effects on the probability of engaging in the productive activities of school or work.[6] A growing body of research focuses specifically on the potentially deleterious effects of out-of-school suspensions, including Adukia, Feigenberg, and Momeni (2023), Fabelo et al. (2011), Shollenberger (2015), and Wolf and Kupchik (2017), although these studies do not produce strong causal evidence of the effect of out-of-school suspensions on future engagement with the criminal justice system, schooling or work.

III. Texas Administrative Data

Extensive administrative data for the large and diverse state of Texas provide a unique opportunity to investigate how principals affect the skill development and longer-term outcomes of their students. The Texas administrative data, housed at the Texas Schools Project at the University of Texas at Dallas, include information on elementary and secondary schooling provided by the Texas Education Agency (TEA), information on post-secondary schooling provided by the Texas Higher Education Coordinating Board (THECB), information on employment provided by the Texas Workforce Commission (TWC), and computerized criminal history (CCH) data provided by the Texas Department of Public Safety. We create matched panel data sets of the universe of students, teachers, and principals.

The sample used in the analysis of post-secondary schooling and employment includes 8th grade cohorts from 2004 to 2011, and the sample used in the analysis of engagement in the criminal justice system includes 8th grade cohorts from 2001 to 2012; the shorter panel used in the analysis of post-secondary schooling and employment results from limited availability of National Student Clearinghouse data that complement the administrative data provided by the THECB.[7] To avoid complications introduced by principal transitions, the analytic samples include only students who begin middle school with a principal who remains in that school through the year of on-time graduation from middle school for that cohort. A strength of the Texas data is the large number of

[6] The focus on a single school district including related limitations in the criminal justice data raise concerns about the influences of nonrandom sample attrition.
[7] While there is some variation, the typical middle school in Texas includes grades 6-8.

principals and schools that are in the sample. For example, the analysis of engagement in the criminal justice system sample includes 1,363 principals in 713 schools.

The Public Education Information Management System (PEIMS), TEA's statewide educational database, reports key demographic data including race, ethnicity, and gender for students and school personnel as well as student eligibility for a subsidized lunch. PEIMS also contains detailed annual information on administrators including position and school. The PEIMS data are merged with information on annual achievement in reading and math, absences, and disciplinary infractions.[8]

Measures of future outcomes of students come from the Texas Higher Education Coordinating Board (post-secondary schooling data), the Texas Workforce Commission (quarterly earnings data), and the Texas Department of Public Safety (criminal history data). The Texas administrative data on post-secondary schooling contain information on enrollment in public and independent two- and four- year colleges in the state of Texas, by year and semester. Because many Texans choose to attend college or university outside of Texas, the Texas higher education administrative data miss many college spells. However, the THECB has matched National Student Clearinghouse (NSC) data with the administrative data and added information on spells at post-secondary institutions outside of Texas and at Texas institutions not covered by the state data for the period 2009-2017. This coverage substantially reduces or even eliminates any attrition bias resulting from outmigration, and we therefore restrict the post-secondary schooling analysis samples to these years. We use information on quarterly earnings to generate measures of employment and attachment to the labor market and classify someone as working in a quarter if they earn at least half the minimum wage for at least four hours per week. Finally, we use the criminal history data (CCH) to produce measures of involvement in the criminal justice system as an adult (at least 18 years old). The CCH

[8] Students in Texas were tested in math and reading each year, taking the Texas Assessment of Academic Skills (offered 1993-2002), the Texas Assessment of Knowledge and Skills (offered 2003-2012), or the State of Texas Assessments of Academic Readiness (offered 2012-2019). Reading and math tests each contain approximately 50 questions, although the number of questions and average percent correctly answered varies across time and grades. We transform all test results into standardized scores with a mean of zero and variance equal to one for each subject, grade, and year. Each test was administered each spring to eligible students enrolled in grades three through eight. Many special education and limited English proficient students are exempted from the tests. In each year roughly 15 percent of students do not take the tests, either because of an exemption or because of repeated absences on testing days.

data contain information on the universe of arrests that have a guilty initial disposition from the time of arrest until the final disposition of the sentence. This includes the initial offense charge, changes to that charge, history of pleas, court verdicts, sentence length, and probation conditions. An arrest with a guilty initial disposition serves as the primary criminal justice outcome, and arrests with a guilty finding not overturned by appeal and incarceration as alternative measures. Note that the CCH data set does not include arrests that do not have a guilty initial disposition, and henceforth we refer to arrests with a guilty initial disposition simply as arrests.

We do not have information on employment or engagement with the criminal justice system outside of Texas.[9] To understand the potential impact of mismeasuring employment and criminal justice histories, we investigate the rate of outmigration of students 13-15 years old and their out-of-state activities by race and high school completion status using the 2000 US Census IPUMS (Table 1). Approximately 9 percent of Texas middle school students in 1995 lived outside of Texas in 2000, with the rate being much higher for those with a high school degree. Black students and those with fewer than 12 years of completed schooling are less likely to live outside Texas five years later.

There are some interesting patterns for 19-year-old males (who were predominantly in 8th grade five years earlier) in the probabilities of school attendance, employment, and living in an institution (Appendix Table a1). Although the 2000 census does not separate prison from other institutions, in earlier censuses that included prisons as a separate category, the vast majority of young men in institutions were incarcerated. Among high school graduates, the share living in an institution is below four percent regardless of location or race, and that share is always lower for those living outside Texas. By comparison, the share in an institution is much higher for those without a high school degree and diverges sharply by race. Black students who leave Texas are more than 50 percent more likely to live in an institution than those who remain in Texas, but

[9] Such mismeasurement also affects other studies of school or educator effects on engagement in the criminal justice system. For example, Deming (2011) primarily uses arrest records for the county in which the school district is located plus information on African American students who are incarcerated in a state prison during a subset of the sample period. Bacher-Hicks, Billings, and Deming (2019) use only the county arrest data used in Deming (2011).

nonblack students who leave Texas are less than half as likely to live in an institution. The very low outmigration rate for those without a high school degree and the steps described below to account for unobserved heterogeneity mitigate any bias introduced by unobserved out-of-state arrests for those not arrested as adults in Texas.

Table 2 presents means of the student characteristics (top panel), skill measures (middle panel), and longer-term outcomes (bottom panel) for all students, for the sample of males used in the analysis of criminal justice, and for the sample of students used in the analysis of post-secondary schooling and work. Almost half of all students are Hispanic, 14 percent are Black, and 36 percent are White, while more than half the students qualify for a subsidized lunch. On average students are absent roughly 8.5 days in a year, the probability of receiving at least one disciplinary infraction approaches 25 percent, and the probability of receiving at least one out-of-school suspension equals 11 percent.[10]

The bottom panel reports outcomes based on activities within 6 years of expected graduation from 8th grade. Slightly more than 10 percent of males are ever arrested or have a guilty outcome as an adult. Restricting crimes to serious misdemeanors and felonies, not surprisingly, lowers the shares ever arrested or with a guilty outcome. Finally, a comparison between the final two columns illustrates the importance of accounting for college attendance outside of Texas: the addition of the NSC information increases the rates of college attendance, persistence, and engagement with college or work by at least 4 percentage points.

IV. Empirical Framework

Identification of principal effects on long-term outcomes requires the separation of principal impacts from student, family, and school factors outside of their control. We first estimate middle school principal value-added to cognitive and noncognitive skills and then relate the estimates of skill value-added to the long-term schooling, employment, and criminal justice outcomes. An alternative two-way fixed effects model

[10] School districts are required to report to the state any disciplinary action taken against a student which results in the removal from any part of their regular classroom program. This broad reporting requirement will differ across districts partly on the basis of district procedures and codes of conduct.

that combines these steps by using school switchers to estimate directly principal effects on post-secondary outcomes suffers from two main issues. First, similar to the argument by Chetty, Friedman, and Rockoff (2014b) in the case of teachers, the direct impacts of the principal may be correlated with time-varying unobserved factors that would confound the estimates of principal effects. By restricting principal effects to specific skill development channels in a two-stage framework, unobservable dynamic influences including social and employment networks, family income, and wealth are less likely to introduce biases. Second, at a practical level, Bartanen and Husain (2022) show that the scarcity of principals who lead multiple schools inhibits efforts to use two-way fixed effects models to identify principal effectiveness relative to others in the same connected network of schools. This scarcity also precludes the inclusion of controls for origin and destination schools or districts, the approach used by Chetty and Hendren (2018a, 2018b).

IV.a. The two-stage empirical model

We separate the description of the estimation of principal effects on skills and the estimation of the relationships between the longer-term outcomes and the value-added to skills. However, the two stages are closely linked through the steps taken to address the various threats to identification of principal effects on longer-term outcomes.

IV.a.1. Principal value-added to skills

Each of the k skills for student i in middle school s with principal p in cohort t (SO^{k}_{ispt}) is specified in Equation 1 as a cubic function of lagged achievement (*f*(ACH)), lagged absences (ABS), and lagged receipt of an out-of-school suspension (DIS), a vector of student characteristics (**X**), a principal-by-school fixed effect (θ_{ps}), a cohort fixed effect (π_{t}), and a random error (ϵ_{ipst}). Note that cohort t is defined as the year a student completes 8th grade in the absence of being retained in or skipping a grade, and the lagged initial conditions are measured in 5th grade for students entering middle school in 6th grade (as shown in the equation) or in 6th grade for students entering middle school in 7th grade. The inclusion of a principal-by-school rather than simply a principal fixed effect permits variation in effectiveness by the quality of the principal-school match.

$$(1) \quad SO^{k}_{ispt} = f(ACH_{i5}) + \alpha^{k}ABS_{i5} + \gamma^{k}DIS_{i5} + \beta^{k}\mathbf{X}_{ist} + \theta^{k}_{ps} + \pi^{k}_{t} + \varepsilon^{k}_{ispt}$$

$$\text{for } k \in (\text{achievement}, \text{absences}, \text{out} - \text{of} - \text{school suspensions})$$

Average math and reading achievement proxies cognitive skills, and both absences and receipt of an out-of-school suspension (or any disciplinary infraction) proxy noncognitive skills. Absences and out-of-school suspensions capture unproductive behaviors at the lower ends of the non-cognitive skills distributions, but they unfortunately provide little information on the variation in broader socio-emotional skills such as conscientiousness and executive functioning. Prior research has not yielded any consistent approach to measuring noncognitive skills, as data availability often determines the proxies. For example, grade point average is sometimes presented as a noncognitive skill proxy, but it is by no means clear that it captures variation only in the noncognitive skill dimension.[11] Regardless, the Texas administrative data do not include course grades, GPA for most of the sample period, or other proxies for the aforementioned noncognitive skills.

In addition to the selection of proxies, the timing of skill measurement presents an important choice. We use 9th grade measures of achievement, absences and out-of-school suspensions following matriculation to high school to capture better the development of skills that persist. Middle school test scores and absences are high-stakes outcome in the Texas accountability system and can reflect efforts that raise outcomes while not having any lasting effects on skills.[12] High schools directly affect skill development and adopt varying rules and disciplinary practices, but we control for high school-by-cohort effects in our preferred specification. We also report estimates using 7th grade achievement, absences, and suspension outcomes as a robustness check.

To obtain estimates of θ^{k}_{ps}, we regress SO^{k}_{ispt} on the controls separately for each cohort, use the coefficients to compute the residual for each student, average the residuals

[11] Jackson (2018), for example, creates a single index for noncognitive skills that includes grade point average (GPA), but this measure undoubtedly also captures variation in cognitive skills, as the acquisition of subject specific knowledge affects examination results and grades and is not captured fully by state standardized mathematics and reading tests.

[12] This is also consistent with Gilraine and Pope (2021) who find that longer-term cognitive skills tend to be much more predictive of future education, economic and social outcomes.

over all students in a school, and take the mean of the school averages calculated over all cohorts (Eq. 2).[13] The value-added estimates are then shrunk by Bayesian shrinkage methods.

$$(2)\ \bar{\hat{\theta}}^{k}_{ps} = \sum_{t \mid ps} \hat{\theta}^{k}_{t}$$

Note that the value-added estimates contain unobserved fixed and time-varying school and neighborhood influences, meaning that their variances do not provide valid measures of the variation in principal effectiveness at raising skills.[14] We take several steps to account for these potentially confounding factors in the second stage.

IV.a.2. Principal effects on post-secondary outcomes

The second stage estimation identifies the channels through which school leaders affect longer-term academic, economic, and criminal justice outcomes. Equation 3 shows the base specification of the relationships between post-secondary outcomes superscripted by o and principal effects on cognitive and noncognitive skills.

$$(3)\ \text{Outcome}^{o}_{ishpt} = f(\text{ACH}_{i5}) + \alpha^{o}\text{ABS}_{i5} + \gamma^{o}\text{DIS}_{i5} + \beta^{o}X_{ist} + \delta^{o}_{abs}\bar{\hat{\theta}}^{ABS}_{pst} + \delta^{o}_{dis}\bar{\hat{\theta}}^{DIS}_{pst} + \delta^{o}_{ach}\bar{\hat{\theta}}^{ACH}_{pst} + \upsilon^{o}_{ishpt}$$

Because many students continue to invest in post-secondary schooling and on-the-job training through their early 20s, our early career panels lead us to focus on outcomes other than earnings. Following Smith and Welch (1987), we classify students as either participating in the productive activities of school or work or in the unproductive activity of crime. Our specific measures include college attendance, college persistence (defined as attendance in three consecutive semesters), engagement in the productive activities of school or work, and involvement in the criminal justice system. Consideration of both attendance and persistence illuminates any differences in skill effects on merely attending college as opposed to succeeding in college as reflected by persistence into the second

[13] As discussed below, we also calculate leave-one-out measures of principal effects and use them in a robustness check that examines the potential confounding effects of random cohort shocks.

[14] Bartanen, Husain, and Liebowitz (2024) offer evidence that controlling for time-varying school factors reduces or eliminates the contribution of principals to the achievement variance. However, in analyses of Texas and Chicago Public Schools that also account for turbulence around principal transitions not addressed in the other study, Laing, et al (2016) and Branch, et al (2020) continue to find sizeable and significant contributions of principals to achievement growth.

year. Each of these outcomes reflects activity within six years of the expected completion of middle school based on progressing one grade per year. Superscripts on the δ parameters reflect the fact that the impacts of cognitive and noncognitive skills may differ across outcomes.

IV.b. Accounting for student and school differences

A major concern is that unmeasured student and school factors affect the first stage estimation of principal skills and are then carried through to the estimation of long run impacts in the second stage. Unobserved student heterogeneity including that introduced by endogenous responses to school quality and both fixed and time-varying school and neighborhood effects on the postsecondary outcomes can bias our estimates of θ_{ps}^{k}. Controls for lagged cognitive and noncognitive skills in both stages account for unobserved heterogeneity at the start of middle school, and the estimation of intent-to-treat effects by assigning students to the middle school initially attended regardless of subsequent school changes accounts for endogenous mobility.

Middle school fixed effects in the second stage account for unobserved middle school, high school and neighborhood differences that are time invariant. These fixed effects capture middle school quality including stable components of the teacher corps, the quality of the high school attended following middle school graduation, neighborhood characteristics, the nature of local job markets, the criminal justice environment, and local amenities. However, these fixed effects do not account for time-varying influences that potentially introduce bias.[15]

We add high school-by-cohort fixed effects based on school attended in 9th grade in the second stage to account generally for time-varying factors including changing school district policies, practices and finances, high school characteristics, local labor market opportunities, social and criminal justice conditions, and other neighborhood

[15] Bartanen, Husain, and Liebowitz (2024), Laing, Rivkin, Schiman, and Ward (2016) and Branch, Hanushek, Rivkin, and Schiman (2020) raise concerns about time-varying school factors, but labor market conditions, college costs, police and judicial practices, and other geographic factors may also change over time.

shocks.[16] Since the substantial majority of middle school students matriculate to a pre-specified local high school, these fixed effects also account for time varying factors common to middle schools that send students to the same high school. Because some students attend a high school other than the one connected structurally to their middle school and because their middle school experience may influence that choice, the inclusion of high school-by-cohort fixed effects does remove this channel of potential principal effects.

Importantly, the high school-by-cohort fixed effects do not control for time-varying shocks to specific middle-school cohorts that may introduce a spurious correlation between estimates of principal effects and longer-term outcomes. In a robustness analysis we substitute leave-one-out estimates of principal value-added in place of the measures averaged over all cohorts to examine the impacts of such shocks. Equation 4 shows the mean value-added for cohort t calculated over all other (t') cohorts that attended school s under principal p. Here $\psi_{t'}$ are the weights for each cohort, and they vary by the number of cohorts before or after cohort t and are equivalent to those obtained from an OLS regression of value-added for principal cohort t on a vector of value-added for all other school cohorts under that principal. This produces leave-one out value-added estimates that allow for drift (Chetty, Friedman, and Rockoff (2014a)). Note that the leave-one-out estimator creates a separate skill measure for each cohort and requires that at least two cohorts of students at a school complete middle school under a given principal. This substantially reduces the effective sample sizes, particularly in models with middle school and high school-by-cohort fixed effects.

$$(4)\ \bar{\hat{\theta}}^{k}_{pst} = \sum_{t' \mid ps} \left[\psi_{t'}\left(\sum_{t' \mid ps} \hat{\theta}^{k}_{t'}\right)\right]$$

V. Principal effects on long-run outcomes

This section presents the results of the empirical analysis of principal effects on educational, employment and criminal-justice outcomes. Following the discussion of the

[16] Because middle school-by-cohort and principal-by-cohort fixed effects would be perfectly correlated, we cannot include middle school-by-cohort fixed effects.

main estimates, we illustrate the sensitivity to the substitution of leave-one-out estimates of skill proxies in place of those aggregated over all cohorts, the substitution of disciplinary infractions in place of out-of-school suspensions and the substitution of VA variables based on skills measured in 7th rather than 9th grade. The section concludes with the consideration of heterogeneous effects by outcome predicted risks and race.

V.a. Variation in principal effects on skills

Tables 3 and 4 report the standard deviations of the skill value-added estimates of (θ_{ps}^{k}) and their correlations with one another. The estimates come from specifications without middle school or high school-by-cohort fixed effects and may therefore capture impacts of other factors beyond principals' control. The standard deviations nevertheless inform the interpretation of the magnitudes of the second stage estimates. Table 3 shows that one standard deviation changes in the three dimensions of principal value-added equals 0.12 standard deviations of achievement, 1.7 days absent (approximately one fifth of the mean), a 0.05 probability of a receiving an out-of-school suspension (roughly one third of the mean), and a 0.07 probability of being cited for a disciplinary infraction (roughly one fourth of the mean).

Principals who are better at raising cognitive skills also tend to be better at improving noncognitive skills (Table 4). The correlations between value-added to achievement on the one hand and value-added to absences and the probability of an out-of-school suspension are -0.37 and -0.19, respectively. Interestingly, substituting the any disciplinary infraction measure in place of the out-of-school suspension measure almost doubles the correlation to -0.33. This suggests that high schools may differ less in the threshold for issuing a disciplinary infraction than in the threshold for imposing an out-of-school suspension, perhaps because some schools handle even severe infractions with in-school punishments.

The correlation of 0.55 between VA to absences and out-of-school suspensions is larger than the correlation between either of these and test score VA, though it remains far below 1. This is consistent with the notion that the two proxies capture different dimensions of principal effectiveness, at least in part.

Finally, the correlations between principal effects on skills measured in 7th and 9th grades for achievement, absences and out-of-school suspensions are all around 0.65, while the 7th to 9th grade correlation for disciplinary infraction receipt is less than 0.5 (Appendix Table a2). In contrast to the findings in Gilraine and Pope (2021) for teacher impacts, there is limited evidence that the correlations between principal effects on cognitive and noncognitive skills are larger when the skills reflect longer-term effects as measured in the future (9th grade) than those measured during middle school. The correlation between principal impacts on achievement and on the probability of receiving a disciplinary infraction are almost twice as large in magnitude for 9th grade (-0.33) as they are for 7th grade (-0.18). But the correlation between test-score value-added and absences value-added equals -0.37 in both grades, and the correlation between test score value-added and impact on out-of-school suspensions is slightly higher in magnitude in 7th grade than in 9th grade (-0.21 v -0.19).

V.b. Male engagement with the criminal justice system

Tables 5 and 6 report the relationships between the probability of a future arrest and value-added to cognitive and noncognitive skills for a series of specifications that sequentially add middle school fixed effects and high school-by-cohort fixed effects to account for fixed and time-varying confounding factors. Standard errors are clustered by school in all specifications.

Table 5 reveals significant effects of principal value-added to noncognitive skills on the probability of an arrest.[17] The pattern for out-of-school suspension VA is particularly striking, where the estimates become larger and remain significant at the 0.01 level following the inclusion of the middle school fixed effects. This highlights the importance of accounting for fixed differences in criminal justice and school disciplinary practices and other school and community influences. Because these specifications do not account for time-varying school and neighborhood factors, the full model adds high school-by-cohort fixed effects. The slightly larger coefficient magnitudes for the full model that controls for time-varying unobservables (Column 3) than for the model without high

[17] The small differences in sample sizes across specifications come mostly from missing information on high school attended. Estimates over common samples are virtually identical to those shown in the tables.

school-by-cohort fixed effects (Column 2) indicate that unobserved school and community factors do not inflate the coefficients.

Table 5 also reveals a significant effect of absence value-added on the probability of an arrest. Again, the pattern differs little across outcomes, but in this case the inclusion of the high school-by-cohort fixed effects reduces the magnitudes and precision of the estimates. Given that many absences result from health conditions outside the control of educators, this pattern is consistent with fixed effects amplifying attenuation bias due to measurement error.

The coefficients from the full model shown in Column 3 indicate that a one standard deviation decrease in out-of-school suspensions VA (0.045) reduces the probability of an arrest by 0.8 percentage points, and a one standard deviation decrease in absence VA (roughly 1.7 fewer days absent) lowers the probability of arrest by 0.5 percentage points These effects translate into roughly 5 percent decreases in the probability of an arrest.

The positive and significant achievement VA coefficients in the full model are unexpected. It is possible that conditional on VA to absences and out-of-school suspensions, a higher achievement VA reflects a greater focus on current achievement at the expense of other considerations. We return to this issue in the heterogeneity analysis below.

Offense severity is a primary determinant of the consequences of engagement with the criminal justice system, and we narrow the focus to more severe crimes by excluding arrests that are not associated with a serious misdemeanor or felony. The estimates reported in Table 6 display a similar pattern to those in Table 5. If anything, the Table 6 coefficients based only on more severe offenses tend to be smaller and less precisely estimated.

Sentencing disposition likely also affects the consequences of engagement with the criminal justice system, and we re-estimate the second stage specifications with two alternative outcome measures. The first reclassifies the indicator for ever arrested to zero for those whose guilty findings are overturned on appeal. This stricter definition of arrests changes the arrest indicator for less than 10 percent of males, and the coefficients on the skill VA measures (not shown) change very little. An indicator for ever incarcerated is

the second alternative measure, and only two percent of students, less than 20 percent of those ever arrested, fall into this category. This very small fraction likely contributes to the imprecision of the skill VA coefficients (not shown), but the magnitudes of the coefficients on out-of-school suspension VA and absence VA translate into similar 5 percent effects for one standard deviation changes in the VA measures.

Finally, as noted above, the estimates from the full model might be affected by school-specific cohort shocks. To assess this, we compare leave-one-out estimates that account for such shocks with the estimates based on all cohorts. Appendix tables a3-a5 show the second-stage estimates based on identical samples are quite similar in terms of significance levels and magnitudes across all long-term outcomes. There are only small differences in effect sizes, and they exhibit no consistent pattern. Thus, there is little evidence that middle school-cohort shocks introduce bias, leading us to focus on the all-cohort estimates here and in the subsequent sections.

V.c. Post-secondary schooling and employment

We turn now to the productive activities of college and work and consider the role of middle school principals in shaping these outcomes. In contrast to the findings for engagement with the criminal justice system, Table 7 shows much stronger effects of achievement value-added on college attendance and persistence. All achievement VA coefficients are significant at the 1 percent level, and the full-model estimates of roughly 0.1 suggest that a one standard deviation increase in achievement VA is associated with one percentage point increases in college attendance and persistence. These effects translate into a 2 percent increase in attendance and 4 percent increase in persistence. By comparison, the fixed-effect coefficients on out-of-school suspension VA are small and insignificant, and those on absence VA are small and insignificant in the attendance models and fluctuate between positive and negative in the persistence models.

Errors in the measurement of outcomes introduced by those who leave a state constitute a common problem in the analysis of state administrative data, with uncertain effects on the estimates. The availability of NSC data on out-of-state college attendance eliminates this problem in our analysis of post-secondary schooling, but we can examine the sensitivity of the estimates to the availability of this information. Comparisons of the

full-model estimates in Column 3 with the estimates in Column 4 where the dependent variable misclassifies those who attend college out-of-state as not attending college, suggests that, at least in this sample, mismeasurement introduced by the absence of information on out-of-state college enrollment would have had little impact on the estimates.

Importantly, the null category in the analysis of post-secondary schooling combines those working but not attending school and those neither attending school nor working. Because skill effects on these two groups may differ substantially, we create a new outcome that divides the sample into those engaged in a productive activity, defined as either attending college or working, and those neither attending college nor working.

Table 8 reports coefficients for the same specifications for weak (top panel) and strong (bottom panel) measures of engagement in productive activities, where weak engagement is defined as school attendance or employment in 6 out of the 8 quarters in the two years following expected high school graduation and strong engagement is defined as either persistence to the third semester of college or employment in all 8 quarters. Similar to the post-secondary schooling specifications, both panels show a strong effect of achievement value-added and little or no evidence that the noncognitive skill effects are strongly related to productive outcomes.

The finding that principal effects on noncognitive skills are only weakly related to post-secondary schooling and employment contrasts with some findings on teacher effects including Jackson (2018). This difference likely emanates in part from the difference in noncognitive skill measures used across the studies. Neither absences nor receipt of an out-of-school suspension captures variation across the full distribution of students in grit, patience, resilience, or other socio-emotional skills that may affect post-secondary schooling. Absences reflect the lower bar of showing up, and out-of-school suspensions provides no information on variation in noncognitive skills for the almost 90 percent of the sample who have not received such a punishment. Therefore, it is not surprising that these proxies, particularly receipt of an out-of-school suspension, are much more strongly related to the probability of engagement with the criminal justice system. By comparison, Jackson (2018) constructs a noncognitive skill index from

multiple factors including grade point average (GPA). Although this index likely captures better variation in noncognitive skills across the distribution, it also incorporates variation in subject knowledge not accounted for by state standardized reading and math tests.

V.d. Disciplinary infraction versus out-of-school suspension

Both behavior and punishment practices determine whether an act results in receipt of a disciplinary infraction and whether the punishment includes an out-of-school suspension. Therefore, middle school principals whose students attend different high schools may have the same value-added to receipt of a disciplinary infraction but different value-added to receipt of an out-of-school suspension because 1) high schools mete out different punishments for the same perceived transgressions; 2) the disciplinary infraction mix is more concentrated among severe infractions for one principal than for another; or 3) some combination of the two. Consequently, the substitution of receipt of a disciplinary infraction in place of an out-of-school suspension does not provide direct evidence on the mediating effect of an out-of-school suspension. However, a finding of a much weaker relationship between engagement with the criminal justice system and disciplinary infraction value-added would be consistent with school suspensions amplifying the negative effect of behavior that triggers school punishment.

Table 9 compares the coefficients on out-of-school suspension value-added taken from Tables 5 through 8 with the coefficients on disciplinary infraction value-added that come from identical specifications that include both middle school and high school-by-cohort fixed effects. Columns 1-2 reveal a striking contrast between the small, insignificant and sometimes negative effect of value-added to a disciplinary infraction on the probability of an arrest and the much larger, highly significant coefficient on value-added to an out-of-school suspension. In contrast, the remaining columns show little difference by measure in the relationship with the schooling and employment outcomes; none of the coefficients show a significant negative relationship with any of the schooling or productivity outcomes.

V.e. Skill measurement grade

A concern that strategic behavior could affect skill measures that are high stakes to the principal leads to measuring skills in 9th grade following matriculation to high

school, but it is informative to examine the sensitivity of the estimates to the skill measurement grade. A comparison of the top and bottom panes of Column 1 in Appendix Table a6 shows that the out-of-school suspension value-added coefficient is much larger and more significant in the arrest specifications that use 9th rather than 7th grade skill measures. This difference provides evidence of the importance of separating effects on behavior from school disciplinary practices through comparisons of students in the same high school and cohort.

The finding for absences contrasts that for out-of-school suspensions, showing much stronger relationships between outcomes and absence value-added based on the measurement of absences in 7th grade. High school absences may provide noisier information on the skills related to responsibility and showing up.

Perhaps the primary concerns relate to test scores, the key high-stakes outcome for middle school principals. Contrary to the belief that contemporaneous scores are more prone to the influences of strategic behaviors such as teaching to the test, the measurement of achievement in 7th grade does not dampen the relationship between productive outcomes and test score VA. Moreover, unexpected positive and significant relationship between test score VA and the probability of engagement with the criminal justice system disappears in specifications based on 7th grade test score measures.

V.f. Heterogeneity by outcome risk and race

We investigate heterogeneity by the predicted outcome risks and also by race. Following Deming (2011), we regress each outcome probability on demographic controls and achievement, absences and an indicator for receipt of a disciplinary infraction in the grade prior to middle school entry. Appendix Table a7 shows that test score, absences and disciplinary infractions in the grade prior to middle school entry are all highly significant predictors of all outcomes, as are gender and race.

For each outcome we divide the samples into the top decile, the top quartile, the top half and the lower half of the predicted outcome distribution and run separate regressions for each group using the full models with middle school and high school by cohort fixed effects. The estimates in the upper left panel of Table 10 show a much stronger relationship between the probability of arrest and value added to out-of-school

suspensions for those with the highest predicted probability of arrest. A coefficient of 0.61 indicates that a one standard deviation higher principal VA to out-of-school suspensions reduces the probability of arrest by 2.7 percentage points; the coefficient of 0.029 for those below the median is small and insignificant. Interestingly, the upper right panel of Table 10 produces a similar ordering for the probability of college attendance, though the estimate effect of out-of-school suspension value added is only significant for the top decile sample. By comparison, the effects of principal value added to achievement on college attendance and persistence (lower left panel) are smallest in the highest decile, indicating that VA to achievement growth has larger effects for those outside the top of the initial achievement distribution. Finally, the unexpected positive relationship between the probability of arrest and test score VA appears for those in the top half of the predicted probability of future arrest distribution. This raises the possibility that the benefits of higher test-score VA in terms of college attendance and persistence may come at some cost to males with a high predicted risk of future arrest and a low predicted risk of attending college.

In terms of race, both post-secondary outcomes and the 9th grade skill measures differ sharply for Black and nonblack middle schoolers in Texas, leading us to consider whether the impacts of principals differ by student race. Looking first at skill measures, the top panel of Table 11 shows that Black students are more than 50 percent more likely to receive an out-of-school suspension in 9th grade than Hispanic students and more than three times as likely than nonblack, nonhispanic students. Although the differences in average absences are much smaller, the 9th grade test score gaps between Black and Hispanic students on the one hand and the remaining students on the other exceeds 0.5 standard deviations. Turning to longer-term outcomes, the bottom panel shows that Black males are 50 percent more likely to be arrested than Hispanic males and more than twice as likely than the remaining male students. Although Black students are slightly more likely to attend and persist in college than Hispanic students, sizeable deficits with other students emerge.

We cannot determine the extent to which the racial gaps in out-of-school suspensions and arrests come from disparate treatment or behavioral differences, but we can investigate the possibility of race differences in the relationships between the

principal value-added measures and outcomes. Because of the small number of Black students in many schools, attempts to estimate separate principal value-added measures by race produced very noisy estimates for Black students (not shown), and we continue to rely on the value-added estimates from the full sample.

Table 12 reveals a large race difference in the effects of out-of-school suspension value-added on the probability of an arrest, but smaller differences in the effects on productive outcomes. The full-model coefficient for out-of-school suspension value-added is more than three times as large for Black than for Hispanic males and is not significant in the case of Hispanics or nonhispanic, nonblack students. The coefficient for Black males indicates that a one standard deviation decrease in out-of-school suspension value-added translates into a roughly 5 percent decrease in the probabilities of arrest. By comparison, the full-model absence value-added coefficient is much larger and more significant for nonblack, nonhispanic students; the coefficient for Hispanic males is roughly half as large and less precisely estimated. Finally, there is little or no evidence that higher achievement value-added significantly reduces engagement in the criminal justice system for any group.

By comparison, Table 12 shows highly significant effects of test-score value added on the probabilities Hispanic students attend and persist in college, smaller and less precisely estimated effects for Black students, and quite small and insignificant coefficients for nonblack, nonhispanic students. Estimates on the probability of engagement in productive activities are also much larger and more significant for Hispanic students; there is little or no evidence of significant effects for the other demographic groups and little evidence of significant effects of either absence or out-of-school suspension value added on this outcome for any of the demographic groups in the full models.

VI. Conclusions

The results highlight the importance of school leadership in the development of the skills that improve longer-term outcomes, though striking differences emerge in the

relative importance of different skill dimensions to different outcomes. The estimates for post-secondary schooling and employment reveal large and highly significant effects of principal value-added to cognitive skills on college attendance, college persistence, and strong engagement in productive activities. At the same time, there are much weaker effects of value-added to noncognitive skills for these outcomes. There is also little evidence of any racial differences in the pattern of principal effects on post-secondary schooling or employment.

In contrast, the estimated effects of principal value-added on the probability of an arrest with a guilty disposition reveal a pattern that is almost diametrically opposed to that for the schooling and employment outcomes: there is little or no evidence that middle school principals affect the probability of engagement in the criminal justice system by raising cognitive skills but strong evidence that they affect it through their impacts on noncognitive skills, especially those related to the probability of an out-of-school suspension. In addition, the relationship between the probability of engagement in the criminal justice system and out-of-school suspension value-added is much larger for Black than for nonblack males and correspondingly larger for those with the highest predicted risk of arrest based on information prior to middle school entry. Although we are not able to disentangle the direct effects of suspensions from the more severe behavioral infractions associated with those suspensions, this finding supports further inquiry into the potentially pernicious effects of out-of-school suspensions on future engagement with the criminal justice system.

References

Adukia, Anjali, Benjamin Feigenberg, and Fatemeh Momeni. 2023. "From Retributive to Restorative: An Alternative Approach to Justice." NBER Working Paper Series No. 31675. Cambridge, MA: National Bureau of Economic Research (September).

Bacher-Hicks, Andrew, Stephen B. Billings, and David J. Deming. 2019. "The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime." NBER Working Paper Series No. 26257. Cambridge, MA: National Bureau of Economic Research.

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and interpretation of teacher value added in research applications." In *Handbook of the Economics of Education, Vol 6*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann: Elsevier: 93-134.

Bartanen, Brendan, and Aliza N. Husain. 2022. "Connected Networks in Principal Value-Added Models." *Economics of Education Review*, 90.

Bartanen, Brendan, Aliza N. Husain, and David D. Liebowitz. 2024. "Rethinking Principal Effects on Student Outcomes." *Journal of Public Economics*, 234.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin. 2012. "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals." NBER Working Paper W17803. Cambridge, MA: National Bureau of Economic Research (January).

Branch, Gregory F., Eric A. Hanushek, Steven G. Rivkin, and Jeffrey C. Schiman. 2020. "How Much Does Leadership Matter? Evidence from Public Schools." unpublished manuscript (July).

Card, David, and Alan B. Krueger. 1992a. "Does school quality matter? Returns to education and the characteristics of public schools in the United States." *Journal of Political Economy* 100, no. 1: 1-40.

Card, David, and Alan B. Krueger. 1992b. "School quality and black-white relative earnings: A direct assessment." *Quarterly Journal of Economics* 107, no. 1 (February): 151-200.

Carrell, Scott E, and James E West. 2010. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy* 118, no. 3 (June): 409-432.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014a. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review* 104, no. 9 (September): 2593-2632.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104, no. 9 (September): 2633-2679.

Chetty, Raj, and Nathaniel Hendren. 2018a. “The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects.” *Quarterly Journal of Economics*, 133(3).

Chetty, Raj, and Nathaniel Hendren. 2018b. “The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates.” *Quarterly Journal of Economics*, 133(3).

Chiang, Hanley, Stephen Lipscomb, and Brian Gill. 2016. "Is School Value Added Indicative of Principal Quality?" *Education Finance and Policy* 11, no. 3 (Summer): 283–309.

Clark, Damon, Paco Martorell, and Jonah Rockoff. 2009. "School Principals and School Performance." CALDER Working Paper 38. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research (CALDER) (December).

Deming, David J. 2011. "Better Schools, Less Crime?" *The Quarterly Journal of Economics* 126, no. 4 (November 1, 2011): 2063-2115.

Deming, David J., Sarah Cohodes, Jennifer Jennings, and Christopher Jencks. 2016. "School accountability, postsecondary attainment, and earnings." *Review of Economics and Statistics* 98, no. 5: 848-862.

Dhuey, Elizabeth, and Justin Smith. 2014. "How important are school principals in the production of student achievement?" *Canadian Journal of Economics* 47, no. 2 (May): 634-663.

Dinerstein, Michael, and Isaac M. Opper. 2022. "Screening with Multitasking: Theory and Empirical Evidence from Teacher Tenure Reform." NBER Working Paper No. 30310. Cambridge, MA: National Bureau of Economic Research (January (revised)).

Elango, Sneha, Jorge Luis García, James J. Heckman, and Andrés Hojman. 2016. "Early Childhood Education." In *Economics of Means-Tested Transfer Programs in the United States, Volume II*, edited by Robert Moffitt. Chicago: University of Chicago Press: 235-297.

Fabelo, Tony, Michael D. Thompson, Martha Plotkin, Dottie Carmichael, Miner P. Marchbanks III, and Eric A. Booth. 2011. *Breaking Schools' Rules: A Statewide Study of How School Discipline Relates to Students' Success and Juvenile Justice Involvement*. New York: Council of State Governments Justice Center (July).

Gilraine, Michael, and Nolan G. Pope. 2021. "Making Teaching Last: Long-Run Value-Added." *National Bureau of Economic Research Working Paper Series* No. 29555.

Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb. 2015. "Using Student Test Scores to Measure Principal Performance." *Educational Evaluation and Policy Analysis* 37, no. 1 (March): 3-28.

Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.

Hochbein, Craig, and Brittany C. Cunningham. 2013. "An Exploratory Analysis of the Longitudinal Impact of Principal Change on Elementary School Achievement." *Journal of School Leadership* 23, no. 1: 64-90.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126, no. 5: 2072-2107.

Jackson, C. Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastián Kiguel. 2020. "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment." *American Economic Review: Insights* 2, no. 4: 491-508.

Jacob, Brian A., and Lars Lefgren. 2008. "Can principals identify effective teachers? Evidence on subjective performance evaluation in education." *Journal of Labor Economics* 26, no. 1 (January): 101-136.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review* 47: 180-195.

Laing, Derek, Steven G. Rivkin, Jeffrey C. Schiman, and Jason Ward. 2016. "Decentralized Governance and the Quality of School Leadership." NBER Wroking Paper No. 22061. Cambridge, MA: National Bureau of Economic Research (March).

Lochner, Lance, and Enrico Moretti. 2004. "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports." *American Economic Review* 94, no. 1 (March): 155-189.

Petek, Nathan, and Nolan G. Pope. 2023. "The Multidimensional Impact of Teachers on Students." *Journal of Political Economy* 131, no. 4 (April): 1057-1107.

Rose, Evan K., Jonathan T. Schellenberg, and Yotam Shem-Tov. 2022. "The Effects of Teacher Quality on Adult Criminal Justice Contact." NBER Working Paper No. 30274. Cambridge, MA: National Bureau of Economic Research (July).

Shollenberger, Tracey L. 2015. "Racial Disparities in School Suspension and Subsequent Outcomes: Evidence from the National Longitudinal Survey of Youth." In *Closing the school disipline gap: Equitable remedies for excessive exclusion*, edited by Daniel J. Losen. New York: Teachers College Press: 31-43.

Smith, James P., and Finis Welch. 1987. "Race and Poverty: A Forty-Year Record." *The American Economic Review* 77, no. 2: 152-158.

Wolf, Kerrin C., and Aaron Kupchik. 2017. "School Suspensions and Adverse Experiences in Adulthood." *Justice Quarterly* 34, no. 3: 407-430.

Table 1. Five-year Outmigration Rates for Texas Middle School Students, by completed years of schooling, gender and race

	Schooling Level		
	Less than high school	Greater than high school	All
1. Blacks and nonblacks			
males and females	0.063	0.101	0.09
	13,057	31,954	45,011
males only	0.062	0.114	0.097
	7,285	15,585	22,870
2. Blacks			
males and females	0.038	0.097	0.08
	1,486	3,772	5,258
males only	0.039	0.107	0.084
	857	1,725	2,582
3. nonblacks			
males and females	0.067	0.102	0.091
	11,571	28,182	39,753
males only	0.065	0.115	0.099
	6,428	13,860	20,288

Note: For each entry, the top line provides the share of 18-20-year-olds who lived in Texas five years earlier and live in a different state in 2000. The second line provides the number of observations in the category. Source: 2000 US Census IPUMS data

Table 2. Mean Student Characteristics, Skill Measures, and Outcomes, by Sample

	All students graduating grade 8 by 2012	Criminal justice system sample	College sample	College sample without NSC data
Student characteristics				
Male	0.513	0.511	0.511	
Black	0.147	0.134	0.134	
Hispanic	0.425	0.463	0.472	
White	0.393	0.370	0.361	
Reduced price lunch eligible	0.533	0.542	0.548	
Special Education	0.126	0.125	0.124	
Skill measures				
Grade 7 test score	0.026	0.052	0.050	
Grade 7 absences	6.742	6.655	6.701	
Grade 7 disciplinary infraction	0.254	0.253	0.261	
Grade 7 out-of-school suspension	0.102	0.100	0.106	
Grade 9 test score	0.082	0.091	0.080	
Grade 9 absences	8.832	8.667	8.730	
Grade 9 disciplinary infraction	0.283	0.279	0.286	
Grade 9 out-of-school suspension	0.118	0.114	0.117	
OUTCOMES				
1. Male engagement with the criminal justice system				
a. All crimes				
Arrested	0.126	0.121	0.119	
b. serious misdemeanor or felony				
Arrested	0.082	0.079	0.078	

	All students graduating grade 8 in 2012	Criminal justice system sample	College sample	College sample without NSC data
2. College and work				
Attend college	0.466	0.493	0.542	0.503
Persist in college for 3 semesters	0.273	0.291	0.349	0.295
persist in college for 3 semesters or work in all 8 quarters	0.385	0.408	0.459	0.409
Work or attend college for at least 6 out of 8 quarters	0.514	0.539	0.588	0.540

Notes: The engagement with the criminal justice system sample includes 8th grade cohorts from 2001 to 2012, and the post-secondary schooling and employment sample includes 8th grade cohorts from 2004 to 2011. Arrest includes only arrests with a guilty initial disposition.

Table 3. Standard Deviations of Estimated Principal Value-added, by Skill Measurement Grade

Outcome and grade	Standard deviation
Grade 9	
test score VA	0.125
absences VA	1.701
disciplinary infraction VA	0.066
out-of-school suspension VA	0.045
Grade 7	
test score VA	0.096
absences VA	0.821
disciplinary infraction VA	0.067
out-of-school suspension VA	0.052

Notes: Value-added measures come from the first stage estimates of Equation 1 separately by skill and year. The specification includes the cubic of lagged achievement, lagged absences and lagged receipt of an out-of-school suspension (or disciplinary infraction), all measured in the year prior to middle school entry, and a vector of student characteristics. Regression residuals are then averaged by school and year and aggregated over principal spells at a school according to Equation 2.

Table 4. Correlations Between Estimated Principal Value-added to Skills

	test score VA	absence VA	disciplinary infraction VA	suspension VA
test score VA	1			
absence VA	-0.373	1		
disciplinary infraction VA	-0.327	0.407	1	
out-of-school suspension VA	-0.190	0.549	0.656	1

Notes: Skills are measured in grade 9. See the Table 3 notes.

Table 5. Estimated Effects of Middle School Principal Skill Value-added on the Probability of an Arrest with a Guilty Initial Disposition (standard errors clustered by school in parentheses)

	(1)	(2)	(3)
middle school fixed effects	no	yes	yes
high school-by-cohort fixed effects	no	no	yes
test score VA	-0.0351***	0.0158	0.0456**
	(0.0111)	(0.0134)	(0.0195)
absence VA	0.0041***	0.0042***	0.0030*
	(0.0009)	(0.0011)	(0.0016)
out-of-school suspension VA	0.1006***	0.1502***	0.1717***
	(0.0337)	(0.0474)	(0.0614)
Observations	487,229	487,227	453,566

Notes: The specification includes the cubic of lagged achievement, lagged absences and lagged receipt of an out-of-school suspension (or disciplinary infraction), all measured in the year prior to middle school entry, a vector of student characteristics and cohort fixed effects. Outcomes are measured over the two years following expected high school graduation. * p<0.1; ** p<0.05; *** p<0.01

Table 6. Estimated Effects of Middle School Principal Skill Value-added on the Probabilities of an Arrest with a Guilty Initial Disposition for a Serious Misdemeanor or Felony (standard errors clustered by school in parentheses)

	(1)	(2)	(3)
middle school fixed effects	no	yes	yes
high school-by-cohort fixed effects	no	no	yes
test score VA	-0.0308***	0.0060	0.0277*
	(0.0075)	(0.0111)	(0.0166)
absence VA	0.0045***	0.0037***	0.0027*
	(0.0007)	(0.0010)	(0.0015)
out-of-school suspension VA	0.0390*	0.0892**	0.1152**
	(0.0235)	(0.0386)	(0.0553)
Observations	487,229	487,227	453,566

Notes: See Table 5 notes. The designation of an offense as a serious misdemeanor or felony is based on the measure of offense severity included in the Texas CCH data. * p<0.1; ** p<0.05; *** p<0.01

Table 7. Estimated Effects of Middle School Principal Skill Value-added on College Attendance and Persistence (standard errors clustered by school in parentheses)

	(1)	(2)	(3)	(4)
middle school fixed effects	no	yes	yes	yes
high school-by-cohort fixed effects	no	no	yes	yes
NSC data included	yes	yes	yes	no
1. College Attendance				
test score VA	0.1431***	0.1639***	0.1099***	0.0910***
	(0.0201)	(0.0220)	(0.0258)	(0.0245)
absence VA	-0.0006	-0.0031	-0.0028	-0.0025
	(0.0023)	(0.0024)	(0.0020)	(0.0019)
out-of-school suspension VA	-0.3031***	-0.0164	-0.0326	0.0031
	(0.0694)	(0.0764)	(0.0745)	(0.0689)
2. College Persistence				
test score VA	0.2324***	0.1316***	0.0941***	0.0863***
	(0.0205)	(0.0224)	(0.0205)	(0.0199)
absence VA	0.0006	-0.0037*	0.0016	0.0023
	(0.0018)	(0.0019)	(0.0019)	(0.0017)
out-of-school suspension VA	-0.1891***	0.0514	-0.0334	-0.0150
	(0.0577)	(0.0628)	(0.0666)	(0.0601)
Observations	685,057	685,051	642,173	642,173

Notes: See Table 5 notes. NSC refers to student data from the National Student Clearinghouse. College persistence is defined as attending three consecutive semesters. * p<0.1; ** p<0.05; *** p<0.01

Table 8. Estimated Effects of Middle School Principal Skill Value-added on the Probability of Engaging in Productive Activities (standard errors clustered by school in parentheses)

	(1)	(2)	(3)	(4)
middle school fixed effects	no	yes	yes	yes
high school-by-cohort fixed effects	no	no	yes	yes
NSC data included	yes	yes	yes	no
1. employed or attending school in 6 out of 8 quarters				
test score VA	0.0775***	0.0866***	0.0514**	0.0433*
	(0.0133)	(0.0256)	(0.0262)	(0.0251)
absence VA	-0.0076***	-0.0057***	-0.0022	-0.0015
	(0.0012)	(0.0020)	(0.0022)	(0.0021)
out-of-school suspension VA	0.0165	-0.1158	-0.0649	-0.0300
	(0.0380)	(0.0801)	(0.0782)	(0.0722)
2. employed 8 of 8 quarters or persisting in college				
test score VA	0.1565***	0.0861***	0.0494**	0.0452**
	(0.0150)	(0.0227)	(0.0230)	(0.0223)
absence VA	-0.0042***	-0.0046**	-0.0006	0.0000
	(0.0013)	(0.0020)	(0.0021)	(0.0019)
out-of-school suspension VA	-0.0309	-0.0251	-0.0202	-0.0019
	(0.0395)	(0.0682)	(0.0709)	(0.0667)
Observations	685,057	685,051	642,173	642,173

Notes: See Table 7 notes. * p<0.1; ** p<0.05; *** p<0.01

Table 9. Estimated Effects of Middle School Principal Skill Value-added for Alternative Noncognitive Skill Proxies on Long Run Outcomes (standard errors clustered by school in parentheses)

	(1)	(2)	(3)	(4)	(5)
	Crime Outcome	College and Productivity Outcomes			
	arrested	college attendance	college persistence	employed or attending college 6 of 8 quarters	employed 8 of 8 quarters or persisting in college
1. All offenses					
out-of-school suspension VA	0.1717***	-0.0326	0.0016	-0.0649	-0.0202
	(0.0614)	(0.0745)	(0.0019)	(0.0782)	(0.0709)
disciplinary infraction VA	-0.0014	0.0510	-0.0285	0.0784*	0.0328
	(0.0373)	(0.0404)	(0.0355)	(0.0421)	(0.0376)
2. Serious misdemeanors and felonies					
out-of-school suspension VA	0.1152**				
	(0.0553)				
disciplinary infraction VA	0.0175				
	(0.0311)				

Notes: See Table 5 notes. All specifications include test score VA, absence VA, middle school and high school-by-cohort fixed effects. Arrest includes only arrests with a guilty initial disposition. Separate regressions are run by noncognitive skill proxy. * p<0.1; ** p<0.05; *** p<0.01

Table 10. Middle School and High School by Cohort Estimated Effects of Middle School Principal Skill Value-added on the Probabilitiy of Arrest, College Attendance, College Persistence, and Engagement in a Productive Activities, by Predicted Risk of the Outcome (standard errors clustered by school in parentheses)

Panel A.	**Probability of Arrest**				**College Attendance**			
Predicted risk of future outcome	(1) Highest 10th percentile	(2) Highest 25th percentile	(3) Above the median	(4) Below the median	(5) Highest 10th percentile	(6) Highest 25th percentile	(7) Above the median	(8) Below the median
test score VA	0.0708	0.0778*	0.0635**	0.0189	0.0958	0.0431	0.0381	0.0550*
	(0.0606)	(0.0405)	(0.0305)	(0.0237)	(0.0746)	(0.0485)	(0.0378)	(0.0321)
absence VA	-0.0084	0.0040	0.0047*	-0.0009	0.0017	0.0015	-0.0034	-0.0019
	(0.0053)	(0.0036)	(0.0024)	(0.0022)	(0.0050)	(0.0034)	(0.0028)	(0.0031)
out of school suspension VA	0.6137***	0.2907**	0.1617*	0.0290	-0.0189	-0.2052	-0.0611	-0.0751
	(0.1719)	(0.1169)	(0.0835)	(0.0812)	(0.2020)	(0.1575)	(0.1108)	(0.1148)

Panel B.	**College Persistence**				**Employed or attending school in 6 of 8 quarters**			
Predicted risk of future outcome	(1) Highest 10th percentile	(2) Highest 25th percentile	(3) Above the median	(4) Below the median	(5) Highest 10th percentile	(6) Highest 25th percentile	(7) Above the median	(8) Below the median
test score VA	0.0341	0.0684**	0.0744***	0.0816**	0.0341	0.0684**	0.0744***	0.0816**
	(0.0310)	(0.0272)	(0.0237)	(0.0366)	(0.0310)	(0.0272)	(0.0237)	(0.0366)
absence VA	-0.0021	-0.0016	-0.0010	0.0020	-0.0021	-0.0016	-0.0010	0.0020
	(0.0027)	(0.0020)	(0.0019)	(0.0033)	(0.0027)	(0.0020)	(0.0019)	(0.0033)
out of school suspension VA	0.0522	0.0088	-0.1065*	0.0676	0.0522	0.0088	-0.1065*	0.0676
	(0.0839)	(0.0726)	(0.0627)	(0.1323)	(0.0839)	(0.0726)	(0.0627)	(0.1323)

Notes: Predicted risk is based on behavior and demographic variables in the year prior to middle school entry. Separate risk predictions are estimated for each outcome. Regressions include middle-school and high school-by-cohort fixed effects. See Table 5, 7 and 8 notes. * p<0.1; ** p<0.05; *** p<0.01

Table 11. Mean Skill Measures and Outcomes, by Student Race

	Black students	Hispanic students	Others
1. Skill measures			
Grade 9 test score	-0.255	-0.135	0.386
Grade 9 absences	9.591	10.195	7.200
Grade 9 disciplinary infraction	0.418	0.334	0.185
Grade 9 out-of-school suspension	0.222	0.138	0.061
2. Outcomes			
Arrested (males)	0.197	0.131	0.085
Attend college	0.531	0.475	0.63
Persist in college for 3 semesters	0.314	0.281	0.443
persist in college for 3 semesters or work in all 8 quarters	0.561	0.549	0.649
Work or attend college for at least 6 out of 8 quarters	0.409	0.41	0.538

Notes: See Table 2 notes.

Table 12. Middle School and High School by Cohort Fixed Effect Estimates of the Effects of Middle School Principal Skill Value-added on the Probabilities of Arrest, College Attendance, College Persistence, and Engagement in Productive Activities, by Student Race (standard errors clustered by school in parentheses)

	(1) Blacks	(2) Hispanics	(3) Other	(4) Blacks	(5) Hispanics	(6) Other
Panel A.	**Probability of an Arrest**			**College Attendance**		
test score VA	0.0029	0.0211	0.0275	0.1017	0.1503***	0.0425
	(0.0661)	(0.0246)	(0.0338)	(0.0670)	(0.0311)	(0.0433)
absence VA	-0.0032	0.0038*	0.0074**	0.0028	-0.0026	0.0009
	(0.0051)	(0.0022)	(0.0037)	(0.0054)	(0.0023)	(0.0046)
out of school	0.3083*	0.0890	0.0098	-0.0758	-0.0632	0.1214
suspension VA	(0.1579)	(0.0875)	(0.1155)	(0.1833)	(0.0908)	(0.1535)
Observations	57,192	212,341	178,054	82,482	307,676	247,331

Panel B.	College Persistence			Employed or Attending School in 6 out of 8 Quarters		
test score VA	0.0856*	0.1102***	0.0410	-0.0133	0.0943***	-0.0022
	(0.0490)	(0.0240)	(0.0473)	(0.0571)	(0.0331)	(0.0451)
absence VA	0.0073	0.0021	0.0019	0.0008	-0.0006	-0.0035
	(0.0055)	(0.0020)	(0.0052)	(0.0050)	(0.0026)	(0.0044)
out of school	-0.2559*	0.0568	0.1009	-0.0127	-0.1119	-0.0945
suspension VA	(0.1463)	(0.0773)	(0.1685)	(0.1446)	(0.0926)	(0.1504)
Observations	82,482	307,676	247,331	82,482	307,676	247,331

Notes: Regressions are estimated separately by student race. Principal skill VA come from first stage models estimated over all students. Regressions include middle-school and high school-by-cohort fixed effects. See Table 5, 7 and 8 notes. * p<0.1; ** p<0.05; *** p<0.01

Appendix A

Appendix Table a1. College Attendance, Employment and Institutionalization of 19-year-old Males Who Lived in Texas 5 years Earlier, by Current State of Residence, High School Completion, and Race (Black vs Nonblack Students)

	Less than high school				Equal to or greater than high school			
	attending school	employed and not attending school	in institution	other	attending school	employed and not attending school	in institution	other
All								
not in Texas	0.205	0.48	0.063	0.26	0.486	0.43	0.013	0.073
in Texas	0.34	0.351	0.082	0.249	0.586	0.28	0.017	0.122
Black students								
not in Texas	0.182	0.091	0.364	0.364	0.41	0.525	0.016	0.049
in Texas	0.373	0.151	0.193	0.354	0.547	0.233	0.037	0.191
nonblack students								
not in Texas	0.207	0.517	0.034	0.25	0.493	0.421	0.013	0.075
in Texas	0.335	0.38	0.066	0.233	0.59	0.285	0.014	0.114

Notes: Figures come from the 2000 US Census IPUMS data

Appendix Table a2. Correlations Between Estimates of Principal Value-added to Test Scores, Absences, Probability of Out-of-school Suspension, and Probability of Disciplinary Infraction, by Skill Measurement Grade

	Grade 7				Grade 9			
value-added to	test scores	absences	disciplinary infractions	out-of-school suspensions	test scores	absences	disciplinary infractions	out-of-school suspensions
grade 7 test scores	1							
grade 7 absences	-0.374	1						
grade 7 disciplinary infractions	-0.178	0.274	1					
grade 7 out-of-school suspensions	-0.264	0.507	0.645	1				
grade 9 test scores	0.691	-0.244	-0.158	-0.165	1			
grade 9 absences	-0.382	0.685	0.31	0.483	-0.373	1		
grade 9 disciplinary infractions	-0.219	0.269	0.492	0.465	-0.327	0.407	1	
grade 9 out-of-school suspensions	-0.211	0.398	0.503	0.668	-0.19	0.549	0.656	1

Note. See Table 3 notes.

Appendix Table a3. Leave-one-out Estimates of Principal Skill Value-added on the Probability a Male is Arrested with a Guilty Disposition (standard errors clustered by school in parentheses)

	All Crimes		Serious misdemeanors and Felonies	
VA construction	all cohorts	leave-one-out	all cohorts	leave-one-out
test score VA	-0.0339***	-0.0390***	-0.0292***	-0.0324***
	(0.012)	(0.013)	(0.008)	-0.0091
absence VA	0.0038***	0.0035***	0.0044***	0.0036***
	(0.0010)	(0.0011)	(0.0008)	(0.0011)
out-of-school suspension VA	0.1095***	0.1082***	0.0429*	0.0593**
	(0.0363)	(0.0317)	(0.0257)	(0.0257)
Observations	451,743	451,743	451,743	451,743

Notes: See Table 5 notes. Regressions with no middle school or high school-by-cohort fixed effects. The sample includes only observations from principal by school spells with multiple cohorts. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table a4. Leave-one-out Estimates of Principal Skill Value-added on Probability of Attending College and Persisting for Three Semesters (standard errors clustered by school in parentheses)

	(1)	(2)
VA construction	all cohorts	leave-one-out
1. College Attendance		
test score VA	0.1385***	0.1500***
	(0.0212)	(0.0226)
absence VA	0.0001	-0.0032
	(0.0025)	(0.0021)
out-of-school suspension VA	-0.3259***	-0.2665***
	(0.0747)	(0.0587)
2. College Persistence		
test score VA	0.2312***	0.2249***
	(0.0217)	(0.0227)
absence VA	0.0009	-0.0023
	(0.0020)	(0.0017)
out-of-school suspension VA	-0.2068***	-0.1810***
	(0.0629)	(0.0513)
Observations	643,667	643,667

Notes: See Table 7 notes. Regressions with no middle school or high school-by-cohort fixed effects. The sample incudes only observations from principal by school spells with multiple cohorts. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table a5. Leave-one-out Estimates of Principal Skill Value-added on Probability of Engaging in Productive Activities (standard errors clustered by school in parentheses)

	(1)	(2)
VA construction	all cohorts	leave-one-out
1. employed or attending school in 6 out of 8 quarters		
test score VA	0.0732***	0.0702***
	(0.0140)	(0.0155)
absence VA	-0.0080***	-0.0067***
	(0.0013)	(0.0016)
out-of-school suspension VA	0.0267	-0.0242
	(0.0409)	(0.0451)
2. employed 8 of 8 quarters or persisting in college		
test score VA	0.1556***	0.1465***
	(0.0159)	(0.0173)
absence VA	-0.0045***	-0.0044***
	(0.0014)	(0.0014)
out-of-school suspension VA	-0.0266	-0.0666
	(0.0431)	(0.0424)
Observations	643,667	643,667

Notes: See Table 7 notes. Regressions with no middle school or high school-by-cohort fixed effects. The sample incudes only observations from principal by school spells with multiple cohorts. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table a6. Estimated Effects of Principal Skill Value-added on the Probabilities of Attending and Persisting in College, Engaging in Productive Activities, and Engaging in the Criminal Justice System, by Skill Measurement Grade (standard errors clustered by school are in parentheses)

	(1)	(2)	(3)	(4)	(5)
	Criminal justice outcome	Schooling and productivity outcomes			
	arrested	college attendance	college persistence	employed or attending college 6 of 8 quarters	employed 8 of 8 quarters or persisting in college
Skills measured in 9th grade					
test score VA	0.0456**	0.1099***	0.0941***	0.0514**	0.0494**
	(0.0195)	(0.0258)	(0.0205)	(0.0262)	(0.0230)
absence VA	0.0030*	-0.0028	0.0016	-0.0022	-0.0006
	(0.0016)	(0.0020)	(0.0019)	(0.0022)	(0.0021)
out-of-school suspension VA	0.1717***	-0.0326	-0.0334	-0.0649	-0.0202
	(0.0614)	(0.0745)	(0.0666)	(0.0782)	(0.0709)
Skills measured in 7th grade					
test score VA	0.0195	0.1250***	0.0981***	0.0924***	0.0711***
	(0.0216)	(0.0277)	(0.0228)	(0.0279)	(0.0263)
absence VA	0.0024	-0.0056*	-0.0040	-0.0045	-0.0062**
	(0.0027)	(0.0033)	(0.0027)	(0.0032)	(0.0030)
out-of-school suspension VA	0.0413	0.0962*	-0.0310	0.0674	0.0237
	(0.0403)	(0.0568)	(0.0427)	(0.0559)	(0.0535)

Notes: See Table 5 notes. Regressions include middle school or high school-by-cohort fixed effects. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table a7. Coefficients from linear probability regressions used to generate predicted risk of future arrest, college attendance, college persistence, and engagement in productive activities

	Arrested	College Attendance	College persistence	employed or attending school in 6 out of 8 quarters	employed 8 of 8 quarters or persisting in college
lag score	-0.0424***	0.1935***	0.1911***	0.1069***	0.1499***
	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0010)
lag score (sq)	-0.0075***	0.0263***	0.0612***	0.0074***	0.0340***
	(0.0008)	(0.0009)	(0.0009)	(0.0010)	(0.0010)
lag score (cube)	-0.0010***	-0.0001	0.0072***	-0.0000	0.0035***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
lagged absences	0.0037***	-0.0089***	-0.0086***	-0.0073***	-0.0084***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
lagged disciplinary infraction	0.1304***	-0.1311***	-0.1165***	-0.0845***	-0.1075***
	(0.0012)	(0.0017)	(0.0016)	(0.0017)	(0.0017)
Male	-	-0.0513***	-0.0579***	-0.0125***	-0.0326***
	-	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Black	0.0680***	0.0447***	0.0079***	0.0069***	-0.0126***
	(0.0015)	(0.0018)	(0.0017)	(0.0018)	(0.0018)
Hispanic	0.0281***	-0.0307***	-0.0456***	-0.0100***	-0.0247***
	(0.0011)	(0.0012)	(0.0012)	(0.0013)	(0.0013)
Asian	-0.0289***	0.0764***	0.1419***	0.0275***	0.0842***
	(0.0028)	(0.0033)	(0.0032)	(0.0033)	(0.0034)
American Indian or Alaskan	0.0059	-0.0641***	-0.0648***	-0.0626***	-0.0732***
	(0.0081)	(0.0094)	(0.0093)	(0.0097)	(0.0099)
Observations	487476	685313	685313	685313	685313