AER: Insights 2020, 2(4): 491–508
https://doi.org/10.1257/aeri.20200029

School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment[†]

By C. Kirabo Jackson, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastián Kiguel*

Using value-added models on data from Chicago Public Schools, we find that high schools impact students' self-reported socioemotional development (SED) by enhancing social well-being and promoting hard work. Conditional on their test score impacts, schools that improve SED in ninth grade reduce school-based arrests and increase high school completion and college going. For most longer-run outcomes, using both SED and test score value added more than doubles the variance of the explained school effect relative to using test score value added alone. Results suggest that high school impacts on SED can be captured using self-report surveys and SED can be fostered by schools to improve longer-run outcomes. (*JEL* I21, J24, K42)

Literature in economics, psychology, and sociology documents that socioemotional skills and mindsets such as adaptability, grit, motivation, empathy, conflict resolution, problem solving, and teamwork are strongly related to education and adult outcomes (Farrington et al. 2012; Duckworth et al. 2007; Dweck 2006; Lindqvist and Vestman 2011; Heckman and Rubinstein 2001; Borghans, Weel, and Weinberg 2008; Waddell 2006; Kautz et al. 2014; Deming 2017). These skills and mindsets (also known as soft or noncognitive) are distinct from the numeracy and literacy skills emphasized in most traditional education systems. In response to this growing knowledge base, many high schools are training teachers to attend to

* Jackson: Northwestern University (email: kirabo-jackson@northwestern.edu); Porter: Mindset Scholars Network (email: shanette@mindsetscholarsnetwork.org); Easton: University of Chicago Consortium on School Research (email: jqeaston@uchicago.edu); Blanchard: University of Chicago Consortium on School Research (email: alyssablanchard@uchicago.edu); Kiguel: Northwestern University (email: skiguel@u.northwestern.edu).
Amy Finkelstein was coeditor for this article. The authors thank the staff at Chicago Public Schools, particularly the Office of Social and Emotional Learning, and the University of Chicago Consortium on School Research for providing access to, and information about, the Chicago Public Schools data. This paper benefited from discussion with seminar participants at the UChicago Consortium, and data management was facilitated by their archivist, Todd Rosenkranz. The authors acknowledge funding for this research from the Bill & Melinda Gates Foundation. The content is solely the responsibility of the authors.
[†]Go to https://doi.org/10.1257/aeri.20200029 to visit the article page for additional materials and author disclosure statement(s).

socioemotional development (SED) and are incorporating socioemotional learning into their curriculums and self-report assessments.[1]

But can high schools influence self-reports of SED, and does it matter for long-run outcomes? Education policy and practice have preempted definitive evidence that SED can be meaningfully shaped in high school. For example, while intervention research has demonstrated that some socioemotional factors are malleable, some debate if this is true for all socioemotional factors (Revelle 2007, Rimfeld et al. 2016, Credé et al. 2017). Additionally, because the self-report measures that are typically used to assess SED in schools are susceptible to response biases, there is uncertainty regarding whether one can measure impacts on SED in ways that are informative for policy (West et al. 2016, Dweck and Yeager 2019). Finally, because most research on SED reflects correlations between socioemotional measures and long-run outcomes, evidence on the extent to which *school-generated* improvements on self-reported SED causally improve subsequent outcomes is limited.[2]

To progress on these issues, we leverage detailed data from Chicago Public Schools (CPS) that link students to schools with self-reported survey measures of SED over time. School value-added models seek to identify schools' causal impacts on student outcomes by comparing end-of-year outcomes across schools while conditioning on lagged outcomes and other covariates. These models have been used extensively and yield results similar to causal estimates from randomized lotteries (Deming 2014, Angrist et al. 2017). Using SED measures in a value-added framework, we (i) estimate schools' impacts on SED in ninth grade, (ii) show that these effects are likely causal, and (iii) determine whether attending a school that improves self-reported SED in ninth grade leads to improved high school outcomes and educational attainment. In existing work, Loeb et al. (2018) and Fricke et al. (2019) examine the variance of school effects on socioemotional growth. However, this is the first study to show that these effects are likely causal and to relate schools' socioemotional impacts to longer-run outcomes.

We find that school effects on SED cluster in two domains: promoting hard work and promoting social well-being. Accordingly, we compute leave-year-out estimates of school value added (Chetty, Friedman, and Rockoff 2014; Jackson 2018) on a hard work index and a social well-being index and also standardized achievement tests. Using these leave-year-out estimates, we explore how attending a school that increases SED in other years (i.e., a high SED value-added school) improves both short- and longer-run outcomes. The standard deviation of estimated school effects on test scores and SED are similar (between 0.06σ and 0.09σ), and these effects are all positively correlated with each other.[3] However, conditional on test score value added, high SED value-added schools improve several high school

[1] In 2004, Illinois was the first state to develop socioeconomic learning (SEL) standards and performance indicators. Since then, at the state level, Kansas, Michigan, Minnesota, New York, North Dakota, Tennessee, and Wisconsin have incorporated measures of SEL into their curriculum (Collaborative for Academic, Social, and Emotional Learnings 2020). There are also many school districts and charter networks in other states that have implemented SEL programs.

[2] While some studies find that interventions *can* improve socioemotional measures in the short run (Sisk et al. 2018; Alan, Boneva, and Ertac 2019; Blackwell, Trzesniewski, and Dweck 2007), there is little evidence of long-run educational attainment impacts. One exception is Dee and Penner (2019), who examine an intervention of which socioemotional training was a component.

[3] These estimated magnitudes are in line with Loeb et al. (2018) and Fricke et al. (2019).

outcomes including school-based arrests and graduation and longer-term outcomes like four-year college going and college persistence. For high school completion and four-year college going, using both SED and test score value added more than *doubles* the variance of the explained school effect relative to using test score value added alone. Looking to mechanisms, schools that improve social well-being have larger effects on attendance and behavioral infractions, while those that promote hard work have larger effects on academics.

To take the estimated effect of value added as reflecting schools' causal impacts requires that, on average, there are no unobserved differences in the determinants of outcomes between students that attend high- and low-value-added schools. As such, we conduct several tests that support a causal interpretation of our results. This work moves beyond showing correlations between SED and long-run outcomes by documenting a wide array of short- and medium-run outcomes that are impacted by attending a school that causally improves SED. This work validates SED value added as capturing school impacts on real skills and traits (as opposed to reporting biases). The analysis presents an important early step toward understanding how schools may influence SED of older adolescents, how it can be measured, and how this can be useful for policy.

I. Data

All data used for this project were collected and housed at the University of Chicago Consortium on School Research (UCCSR 2020). We use administrative data from CPS. CPS is a large urban school district with 133 public (neighborhood/charter/vocational/magnet) high schools. CPS students are primarily Black (42 percent) and Latinx (44 percent) and from families with disadvantaged economic backgrounds (86 percent). The main analysis dataset includes cohorts of ninth grade students who attended one of these schools between 2011 and 2017 ($n = 157{,}630$). For longer-run outcomes, we focus on cohorts of ninth grade students between 2011 and 2014 ($n = 55{,}560$) because these students are old enough to have attended college. Only first time ninth graders are included to eliminate sample selection biases due to grade repetition.

Measures.—Our key variables are survey measures of SED:[4] interpersonal skills, school connectedness, academic engagement, grit, and academic effort. Responses are collected by CPS on a survey administered to students in 2008–2009 and then every year from 2010–2011 onward. While schools receive school-level information about these measures, they are not part of Chicago's accountability system. Survey response rates were high (78 percent); however, nonresponse was higher for low achievers (online Appendix Table S3). Note that our analysis of impacts on longer-run outcomes is based on all students irrespective of survey completion. Each survey measure was comprised of several items, and students responded to each item using point scales to indicate agreement (e.g., $1 =$ strongly disagree to $4 =$ strongly agree). Rasch analysis was used to model responses and calculate

[4] These measures were developed by the University of Chicago Consortium on School Research.

a score for each student on each construct (for measure properties, see online Appendix Table S4).

Two of the SED survey measures relate to one's relationship with others in the school. The first is interpersonal skills, which includes the following: "I can always find a way to help people end arguments;" "I listen carefully to what other people say to me;" "I'm good at working with other students;" "I'm good at helping other people." The second such construct is school connectedness, which includes the following: "I feel like a real part of my school;" "people here notice when I'm good at something;" "other students in my school take my opinions seriously;" "people at this school are friendly to me;" "I'm included in lots of activities at school."

The other three SED survey measures capture students' orientation toward hard work. The first of these is academic effort, which includes the following: "I always study for tests;" "I set aside time to do my homework and study;" "I try to do well on my schoolwork even when it isn't interesting to me;" "if I need to study, I don't go out with my friends." The second construct is the perseverance facet of grit, which includes the following: "I finish whatever I begin;" "I am a hard worker;" "I continue steadily towards my goals;" "I don't give up easily." The third construct is academic engagement, which includes the following: "the topics we are studying are interesting and challenging;" "I usually look forward to this class;" "I work hard to do my best in this class;" "sometimes I get so interested in my work I don't want to stop."

We combine the social-related questions into a social index and the hard-work-related questions into a work hard index. The construction of these indices was informed by conceptual frameworks for SEL. Online Appendix A shows that school effects on individual survey constructs cluster into these two broader categories so that this categorization, in addition to being theory driven, is justified by the data. To create each index, we standardize each construct, compute the average of the included measures, and then standardize the index to be mean zero unit variance.

Test Scores: The "hard" skills measure in our data are standardized test scores.[5] To allow for comparability across grades, test scores were standardized to be mean zero unit variance within grade and year among all CPS test takers. For each student we average the standardized math and English scores and then standardize the test score index to be mean zero unit variance.

We collect intermediate outcomes in ninth grade from the administrative CPS database. These include the number of excused or unexcused absences during ninth grade and the number of disciplinary incidents eligible for suspension. We also include an "on track" indicator that identifies students as on track if they earn at least five full-year course credits and no more than one semester F in a core course in their first year of high school. This is used by CPS and is a more accurate predictor of graduation than test scores or demographics (Allensworth and Easton 2005).

A key outcome is having a school-related arrest (among those old enough to have graduated high school). These are arrests for activities conducted on school

[5] Sixth through eighth grade CPS students took the ISTAT prior to 2014 and the NWEA or PARCC thereafter. Ninth graders took the EXPLORE assessments before 2014 and the PARCC thereafter. When ninth grade spring tests are not administered, we use the tenth grade fall test. This affects 4 percent of the sample and does not change the results.

grounds, during off-campus school activities, or due to a referral by a school official. During our sample period, 4.06 percent of students had a school-based arrest, 5.3 percent of males, and 7.9 percent of Black males. Roughly 20 percent of juvenile arrests in 2010 were school-based arrests (Kaba and Edwards 2012), so these have important long-term implications. Our first longer-run educational outcome is high school completion. About 79 percent of first time ninth graders in CPS graduate high school. Our second key long-run outcome is enrolling in college. Our college data come from the National Student Clearinghouse (NSC 2020) and are merged with all CPS graduates and accessed through the University of Chicago Consortium on School Research. We code a student as enrolling in college if they are observed in the NSC data within two years of expected high school graduation (2011 through 2014 cohorts only). About 53 percent of first time ninth graders enrolled in college. The data are summarized in Table 1.

II. Methods

Our analysis involves two key steps. First, we aim to identify schools that improve students' SED and test scores. With this information, we then estimate the effects of attending schools that improve these measures. We discuss each step in turn.

Step 1: Identifying School Impacts on SED and Test Scores.—We use value-added models to estimate schools' impacts on ninth grade SED and test scores. Our value-added model seeks to isolate the causal effects of individual schools on student measure $q \in Q = \{\text{test scores, social well-being, hard work}\}$ by comparing measures at the end of ninth grade to those of similar students (with the same incoming test scores, survey measures, course grades, discipline, attendance, and demographics, all at the end of eighth grade) at other schools. A school's value added on a measure q captures how much that school increases that measure between eighth and ninth grade relative to the observed changes for similar students (based on the attributes listed above) at other schools. Formally, we model the ninth grade measure q of student i who attends school j with characteristics Z_{ijt} in year t as below. The term Z_{ijt} includes lagged measures (i.e., eighth grade test scores, surveys, discipline, and attendance), gender, ethnicity, and free-lunch status. We include school-level averages of all individual lagged outcomes. We also include the socioeconomic status (SES) of the student census block proxied by average occupation status and education levels. For each measure q, to obtain estimates of the impacts of attending school j in year t relative to the average school (i.e., $\theta_{jt,q}^{VA}$), we estimate (1) below, where $\upsilon_{ijt,q} = \theta_{j,q}^{VA} + \varepsilon_{ijt,q}$:

$$q_{ijt} = \beta_q Z_{ijt} + \upsilon_{ijt,q}, \tag{1}$$

where $\upsilon_{ijt,q}$ is the true student-level error from (1), $u_{ijt,q}$ is the empirical student-level residual obtained after estimation of (1). The average school-year-level residuals from this regression is our estimated impact on measure q of attending a school in a given year. Where N_{jt} is the number of students attending school j in year t, this is

$$\hat{\theta}_{jt,q}^{VA} = \sum_{i \in jt} (u_{ijt,q}) / N_{jt}. \tag{2}$$

Table 1—Summary Statistics

	Analytic sample: short term (2011–2017)		Analytic sample: long term (2011–2014)	
	Mean	SD	Mean	SD
Demographics				
Female	0.502	0.500	0.500	0.500
Special education (individualized education program)	0.182	0.386	0.169	0.375
Free lunch	0.777	0.416	0.769	0.422
Reduced-price lunch	0.075	0.264	0.086	0.280
Census block SES	−0.450	0.874	−0.442	0.873
White	0.0898	0.286	0.0876	0.283
Black	0.403	0.490	0.427	0.495
Native American	0.0017	0.041	0.0018	0.043
Asian/Pacific Islander	0.0327	0.178	0.0327	0.178
Latinx	0.460	0.498	0.440	0.496
Ninth grade intermediate outcomes				
Days absent in ninth grade	14.999	18.633	17.584	21.194
GPA in ninth grade	2.426	1.007	2.264	1.043
Diciplinary incidents in ninth grade	0.0769	0.419	0.106	0.452
On track in ninth grade	0.849	0.358	0.799	0.400
Long-term outcomes				
Any school-based arrest			0.0406	0.197
Graduation			0.743	0.437
Enrolled in any college within two years			0.532	0.499
Enrolled in a two-year college within two years			0.276	0.447
Enrolled in a four-year college within two years			0.343	0.475
Observations	157,630		55,560	

Notes: Number of observations may vary by variable due to missingness and variation in cohorts for which a variable was collected. For more information, see online Appendix Table S3.

If unobserved determinants of student outcomes are unrelated to our value-added estimates, $\hat{\theta}_{jt,q}^{VA}$ will an be unbiased estimate of the value added on school j in year t for measure q.

When using value added to *predict* outcomes for a particular cohort, we exclude data for that same cohort when estimating value added to avoid mechanical correlation. As in Jackson (2014), these leave-year-out (or out-of-sample) predictions of school effectiveness are based on the value added for the same school in *other years*. If the value added in year $t+1$ were equally predictive of outcomes in year t as those in $t+4$ or any other year, then the best leave-year-out predictor for a school would be the average value added for that school *in all other years*. However, adjacent years tend to be more highly correlated with one another than less temporally proximate years (see online Appendix B). Accordingly, following Chetty, Friedman, and Rockoff (2014) to improve precision, we place more weight on value added for years that are more highly correlated with the prediction year. This puts more weight on adjacent years and less weight on years that are temporally farther away. Our leave-year-out predictor for measure q in year t is

$$\hat{\mu}_{jt,q} = \sum_{m=t-l}^{t-1} \hat{\psi}_{m,q}\left[\hat{\theta}_{jm,q}^{VA}\right]. \tag{3}$$

The vector of weights $\hat{\psi}_q = \left(\hat{\psi}_{t-l,q}, \ldots, \hat{\psi}_{t-1,q}, \hat{\psi}_{t+1,q}, \ldots, \hat{\psi}_{t+l,q}\right)'$ are selected to minimize mean squared forecast errors (Chetty, Friedman, and Rockoff 2014). A school's predicted value added on measure q is our best prediction *based on other years* of how much that school will increase measure q between eighth and ninth grade relative to the improvements of similar students at other schools. We use leave-year-out predictions for all analyses, but for brevity, refer to them simply as value added.

Step 2: Estimating the Effect of Value Added on Outcomes.—To quantify the effect of attending a school with one standard deviation higher predicted value added on outcomes, we regress each outcome on the standardized predicted value added for the different indexes (plus controls). Specifically, where Y_{ijt} is an outcome and $\hat{\mu}_{jt,q}$ is the standardized out-of-sample predicted value added on measure $q \in Q = \{$test scores, social well-being, hard work$\}$, we estimate the following model by OLS:

$$Y_{ijt} = \sum_{q \in Q} \beta_q \hat{\mu}_{jt,q} + \beta_1 Z_{ijt} + \tau_t + \varepsilon_{ijt}. \tag{4}$$

All variables are as defined above, and τ_t is a year fixed effect. Standard errors are adjusted for clustering at the school level.[6]

In some models we report estimates using only a single value added predictor, while in others we include several at once. For each model, we report the standard deviation of the predicted school effects (based on value added) on the respective outcome. We compute this in two steps. First, we compute the predicted effect by estimating equation (4) and computing $\hat{F}_{j,q} = \sum_{q \in Q} \hat{\beta}_q \hat{\mu}_{jt,q}$ for each school. This is the predicted impact of attending school j on outcome q based on the value added for school j and the linear relationship between the value-added estimates and the outcome. Next, we compute the sample standard deviation of this predicted effect $\hat{\sigma}_{\hat{F}}$. This is the standard deviation of the predictable impact of schools based on the value-added estimates. With a single value-added measure $\hat{\sigma}_{\hat{F}} \approx \hat{\mu}_{jt,q}$ because the standard deviation of $\hat{\mu}_{jt,q} \approx 1$ in the estimation sample. *They are not exactly the same because the value addeds are standardized for the population rather than the estimation sample.* By comparing the explained school-level variability in models that include only test score value added, only SED value added, and all the value addeds, we can assess how much additional predictive power there is in each value-added measure over the others.

Taking the estimated effect of value added as reflecting schools' causal impacts requires that, on average, there are no unobserved differences in the determinants of outcomes between students that attend high- and low-value-added schools. We assess this in online Appendix C, where we show that value added is unrelated to observable determinants of student outcomes, validate our estimates using quasi-random variation based on school attendance zones, and show that our estimates are similar

[6] Individuals with missing eighth grade surveys or test scores are given imputed values. We regress each survey measure or test score on all observed pre-eighth grade covariates. We then obtain predicted eighth grade values based on these regressions and replace missing values with the predictions. Results are similar with and without imputation.

using within-family variation. These tests support a causal interpretation of our results.

III. The Impact of School Value Added on SED and Test Scores

Here we establish that schools' SED value addeds do in fact predict school impacts on SED. The coefficients in the top panel of Table 2 represent the effect of attending a school with one-standard deviation higher value added (for each measure) on self-reported ninth grade social well-being. For brevity, we will refer to social well-being value added as social value added. As expected, social value added is highly predictive of school impacts on social well-being. Using only social value added, the coefficient of 0.0904 (p-value < 0.01) in column 1 indicates that attending a school that has one standard deviation higher predicted social value added (i.e., going to a school at the eighty-fifth percentile of the social value-added distribution versus one at the median) would improve social well-being by 9.04 percent of a standard deviation—compelling evidence that schools can, and do, impact reported social well-being and that these impacts are persistent over time.

While work hard value added predicts ninth grade social well-being on its own, in models that include both SED value addeds (column 4), work hard value added has little additional explanatory power. Indeed, the standard deviation of the predictable school effect is largely the same using both SED value addeds as just using social. In column 3, we explore how school impacts on test scores predict social well-being. In models with test score value added only, the coefficient on test score value added is 0.0380. This is much smaller than the predictive power of the social well-being value added. Indeed, relative to using the SED value addeds, adding test score value added increases the explained standard deviation by $100 \times [(0.0908/0.0898) - 1] = 1.11$ percent. The explained variance is the square of the explained standard deviation, so this is an increase in the explained variance of $100 \times [(0.0908/0.0898)^2 - 1] = 2.24$ percent. That is, virtually all of the detectable variation in school impacts on self-reported social well-being (using all three value addeds) is captured by social value added.

We now turn to school impacts on the self-reported work hard dimension in the middle panel of Table 2. In models with work hard value added only, attending a school that has one standard deviation higher work hard value added improves self-reported work hard in ninth grade by 6.39 percent of a standard deviation (p-value < 0.01). Models that use social value added only are similar to those that use work hard value added only. In models that include school value added on both SED measures simultaneously, the coefficient on work hard is the largest (0.0455), but that for social well-being is statistically significant (0.0261). Relative to work hard value added alone, adding social value added increases the explained school-level standard deviation and variance by a modest 4.38 and 8.97 percent, respectively. In models with test score value added only (column 3), the coefficient on test score value added is 0.0294 (p-value < 0.01). This is much smaller than the predictive power of work hard value added. In models including all three value-added measures (column 5), test score value added has little independent explanatory power. Relative to using the SED value addeds, adding test score value added increases the explained variance by only 1.81 percent. The

Table 2—Effects on Ninth Grade Measures

	(1)	(2)	(3)	(4)	(5)
Outcome = social well-being in ninth grade (124,685 observations)					
Social value added	0.0904			0.0939	0.0912
	(0.00799)			(0.00814)	(0.00828)
Work hard value added		0.0620		−0.00452	−0.00708
		(0.00911)		(0.0109)	(0.0110)
Test scores value added			0.0380		0.0131
			(0.00592)		(0.00574)
p-value from F-test of work hard and social				0.0000	0.0000
Standard deviation of predicted school effect	0.0897	0.0619	0.0424	0.0898	0.0908
Outcome = work hard in ninth grade (124,487 observations)					
Social value added	0.0606			0.0261	0.0243
	(0.00792)			(0.00699)	(0.00708)
Work hard value added		0.0639		0.0455	0.0438
		(0.00888)		(0.0117)	(0.0120)
Test scores value added			0.0294		0.00861
			(0.00568)		(0.00532)
p-value from F-test of work hard and social				0.0000	0.0000
Standard deviation of predicted school effect	0.0601	0.0638	0.0328	0.0666	0.0672
Outcome = test scores in ninth grade (102,235 observations)					
Social value added	0.0600			0.0385	0.0315
	(0.0108)			(0.0122)	(0.0102)
Work hard value added		0.0573		0.0291	0.0213
		(0.0114)		(0.0143)	(0.0125)
Test scores value added			0.0612		0.0475
			(0.0111)		(0.0108)
p-value from F-test of work hard and social				0.0000	0.0000
Standard deviation of predicted school effect	0.0595	0.0572	0.0683	0.0626	0.0830

Notes: Results are based on regression of ninth grade measures on out-of-sample social, work hard, and test score value added. All models include individual demographic controls (race/ethnicity, free and reduced-price lunch, and gender), eighth grade lags (math and English language arts test scores, survey measures, absences, and discipline), and school-level averages for all the demographics and lagged measures as well as year fixed effects. We also include the SES of the student census block proxied by average occupation status and education levels. For each model, we estimate the standard deviation of the predicted school impacts based on the value-added measures included in the model (i.e., $\hat{\sigma}_{\hat{F}}$). In columns 4 and 5, we report the p-value associated with the test that the coefficients on work hard value added and social value added are jointly equal to zero. Missing eighth grade measures were imputed using seventh grade measures and demographic characteristics. Sample size varies by outcome due to missingness. Results do not change if we restrict results to a balanced sample. Robust standard errors adjusted for clustering at the school level in parentheses.

best predictor of a school's impact on work hard is school value added on work hard—social well-being value added has a small amount of independent predictive power for impacts on the work hard dimension, while test score value added has no independent predictive power.

We conduct similar analyses for ninth grade test scores (lower panel of Table 2). In models that use test score value added only (column 3), attending a school with one standard deviation higher predicted test score value added increases ninth grade test scores by 6.12 percent of a standard deviation (p-value < 0.01).[7] Interestingly, value added on SED measures are almost as good predictors of

[7] The magnitudes of these estimates are in line with Jackson (2013) based on North Carolina schools.

impacts on test scores as test score value added. In models with both the SED and test score value addeds, each measure independently predicts test scores in ninth grade. Relative to using test score value added only, adding the SED value addeds increase the explained variance by 47 percent. This stands in stark contrast to the pattern for SED measures, where the vast majority of a school's effect on SED is captured by the SED value addeds. Remarkably, value added on SED contains considerable independent explanatory power in explaining school impacts on test scores—suggesting that SED may be foundational for academic success.

While test score and SED value added predict independent variation in test scores and SED, the SED and test score value addeds have correlation of about 0.4 (see scatterplots in Figure S1). However, from a policy perspective the key question is whether the independent variation in SED value added predicts improved longer-run outcomes. We explore this below.

IV. Impacts on Educational Attainment

The first educational attainment outcome we examine is high school graduation in Table 3. In models that include a single value-added measure at a time, a one standard deviation increase in test score value added (i.e., going from an average school to one at the eighty-fifth percentile of the effectiveness distribution) increases the likelihood of high school graduation by about 1 percentage point (column 3), that for social value added is 1.56 percentage points (column 1), and that for hard work value added is similar at 1.59 percentage points (column 2). We visually present the nonparametric relationships between the value added and high school graduation in Figure 1. The SED value addeds are each stronger predictors of school impacts on high school graduation than test score value added. In the model with all three value-added measures (column 5), the standard deviation of the predicted school impacts is 0.0187 compared to only 0.0116 for test score value added only. That is, using the SED value added increases the explained variance by about 160 percent relative to using test score value added alone. The p-value on the joint significance of the two SED value-added measures, conditional on test score value added, is less than 1 percent. Because much of the predictable school impacts on high school completion are captured by the surveys, using the test score value added increases the explained variance by only 19.5 percent relative to using only the two SED value addeds.

The second panel of Table 3 reports the impact on enrolling in college within two years of expected high school graduation. In models with a single value-added measure, a one standard deviation increase in test score value added increases college going by 1.5 percentage points, that for social value added is 1.72 percentage points, and that for hard work value added is 1.96 percentage points. The standard deviation of the predicted school impacts using all three value-added measures is 2.28 percentage points. Using the SED value addeds increases the explained variance by 86 percent relative to using test score value added alone. In contrast, using the test score and SED value addeds increases the explained variance by only 27 percent relative to using only the two SED value addeds. In sum, relative to using test score value added, the SED value addeds increases our ability to identify school impacts on longer-run outcomes considerably.

Table 3—Effects on Longer-Run Educational Attainment

	(1)	(2)	(3)	(4)	(5)
Graduate high school (82,146 observations)					
Social value added	0.0156			0.00892	0.00774
	(0.00356)			(0.00439)	(0.00443)
Work hard value added		0.0159		0.00951	0.00837
		(0.00372)		(0.00479)	(0.00496)
Test scores value added			0.0104		0.00668
			(0.00242)		(0.00240)
p-value from F-test of work hard and social				0.0000	0.0007
Standard deviation of predicted school effect	0.0155	0.0159	0.0116	0.0171	0.0187
Enroll in college within two years of expected high school graduation (55,560 observations)					
Social value added	0.0172			0.00735	0.00606
	(0.00520)			(0.00449)	(0.00452)
Work hard value added		0.0196		0.0143	0.0113
		(0.00588)		(0.00625)	(0.00567)
Test scores value added			0.0150		0.0105
			(0.00455)		(0.00377)
p-value from F-test of work hard and social				0.0028	0.0070
Standard deviation of predicted school effect	0.0171	0.0196	0.0167	0.0202	0.0228
Enroll in four-year college within two years of expected high school graduation (55,560 observations)					
Social value added	0.0296			0.0151	0.0135
	(0.00809)			(0.00691)	(0.00709)
Work hard value added		0.0320		0.0212	0.0174
		(0.00925)		(0.0101)	(0.00953)
Test scores value added			0.0207		0.0128
			(0.00579)		(0.00417)
p-value from F-test of work hard and social				0.0007	0.0020
Standard deviation of predicted school effect	0.0294	0.0320	0.0231	0.0336	0.0360
Persist in college within three years of expected high school graduation (55,560 observations)					
Social value added	0.0132			0.00564	0.00450
	(0.00432)			(0.00408)	(0.00412)
Work hard value added		0.0151		0.0110	0.00837
		(0.00491)		(0.00548)	(0.00506)
Test scores value added			0.0125		0.00919
			(0.00418)		(0.00356)
p-value from F-test of work hard and social				0.0060	0.0147
Standard deviation of predicted school effect	0.0131	0.0151	0.0140	0.0156	0.0181

Notes: Results are based on regression of educational attainment on out-of-sample social, work hard, and test scores value added. All models include individual demographic controls (race/ethnicity, free and reduced-price lunch, and gender), eighth grade lags (math and English language arts test scores, survey measures, absences, and discipline), and school-level averages for all the demographics and lagged measures as well as year fixed effects. We also include the SES of the student census block proxied by average occupation status and education levels. For each model, we estimate the standard deviation of the predicted school impacts based on the value-added measures included in the model (i.e., $\hat{\sigma}_{\hat{F}}$). In columns 4 and 5, we report the p-value associated with the test that the coefficients on work hard value added and social value added are jointly equal to zero. Missing eighth grade measures were imputed using seventh grade measures and demographic characteristics. Robust standard errors adjusted for clustering at the school level in parentheses.

To delve deeper into the college results, we explore impacts on attending a two-year or four-year college. We find no effects on two-year college going (not shown).[8] However, we find large effects on four-year college going (third panel of

[8] Better schools may cause some students to shift from two- to four-year colleges. As such, a null effect overall may reflect such shifting for some students and increased two-year college going for others. We explore this in future work.

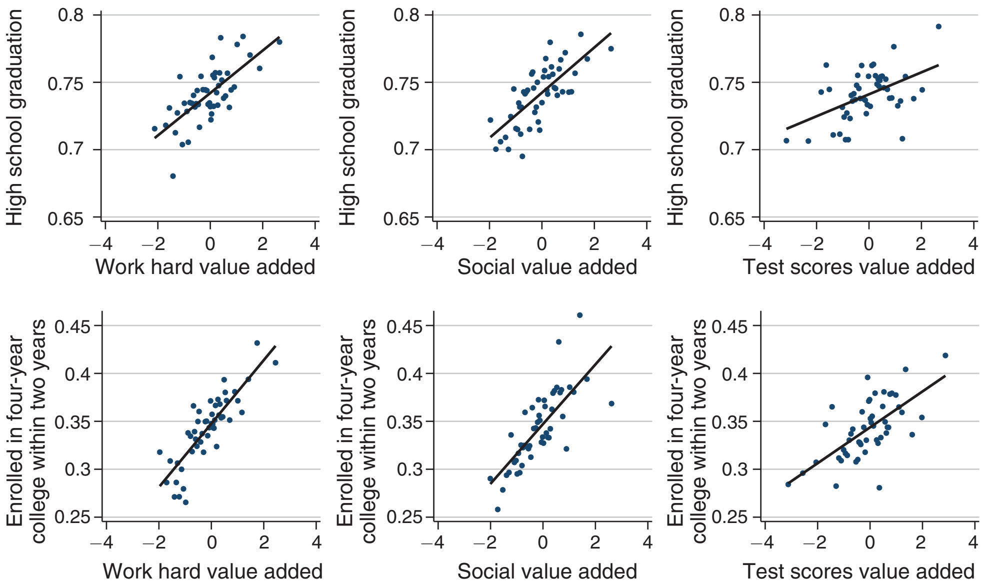

Figure 1. Binned Scatterplot of High School Graduation and Four-Year College Going by Value Added

Notes: This figure presents binned scatterplots of high school graduation (top) and four-year college going (bottom) against each value-added measure conditional on all controls. The controls include individual demographic controls (race/ethnicity, free and reduced-price lunch, and gender), eighth grade lags (math and English language arts test scores, survey measures, absences, and discipline), and school-level averages for all the demographics and lagged measures as well as year fixed effects. We also include the SES of the student census block proxied by average occupation status and education levels. There are 50 bins in each plot.

Table 3). In the single value-added models, a one standard deviation increase in test score value added increases four-year college going by 2.07 percentage points, that for social value added is larger at 2.96 percentage points, and that for work hard value added is largest at 3.2 percentage points. The nonparametric binned scatterplots are in Figure 1. A one standard deviation increase in predicted school impacts (based on all three value-added measures) would increase four-year college going by 3.6 percentage points. Using the SED value addeds increases the explained variance by over 140 percent relative to using test score value added alone. In contrast, using all three measures versus using the SED value addeds increases the explained variance by only 15 percent—reinforcing the relative importance of school impacts on SED.

A comparison of the high school completion and college-going effects sheds light on how many marginal college goers were nonmarginal high school graduates. If all marginal high school graduates went to college, it could explain *at most* $0.0187/0.0228 = 82$ percent of the overall college-going effect and $0.0187/0.036 = 52$ percent of the four-year college going effect. This suggests that the increased college going is among both marginal and nonmarginal high school graduates.

Finally, the bottom panel of Table 3 shows estimated impacts on persisting in college beyond the first year. Each value-added estimate predicts positive impacts on being observed in a second year of college. The standard deviation of the predicted school impacts (using value added on all three measures) on persistence is 1.81 percentage points. The same figure for college going is 2.28 percentage points, suggesting a persistence rate of roughly $1.81/2.28 = 0.79$. This is slightly *higher* than the national average persistence rate of about 0.72.

Potential Mechanisms.—To shed light on mechanisms, we estimate effects on being on track, discipline, and attendance (see Table 4). Because attendance data are audited, and we focus on offenses (as opposed to actual suspensions), changes in these measures likely reflect behavioral change. However, we cannot rule out that schools that differ on their value added also differ in how they report discipline, assign grades, or measure attendance. The first of these is being on track—an indicator of academic performance (top panel). Students who are on track in Chicago at the end of ninth grade are more than three times more likely to graduate high school in four years than off-track students—so this is a good potential mediator. Each value-added measure individually predicts being on track (columns 1 through 3). However, the effects are larger for the SED measures. Attending a school with one standard deviation higher social, work hard, or test score value added leads to a 1.95, 2.1, and 1.41 percentage point increase in the likelihood of being on track, respectively. Relative to a model with test score value added only, the explained variance using both test score and SED value added is over 160 percent larger, indicating that (i) much of what schools may do to keep students on track to graduate high school is largely unmeasured by impacts on standardized tests and (ii) school impacts on self-reported survey measures capture much more of a school's impact on staying on track than impacts on test scores.[9]

Impacts on ninth grade absences are in the second panel of Table 4. The point estimates in columns 1 through 3 indicate that each value-added measure individually predicts better attendance in ninth grade. However, social value added explains more variance than the other two. A one standard deviation increase in test score value added leads to 0.772 fewer absences. By comparison, a one standard deviation increase in social value added reduces absences by 1.3 days—an effect size of roughly 0.07σ or an 8.6 percent reduction compared to the average. Relative to the test score value added only model, adding the SED value addeds increases the explained variance by over 150 percent. The fact that social value added is the most predictive of reduced absences is consistent with work finding that students who feel a greater sense of belonging are more likely to attend school (Walton and Brady 2017).

Next we examine impacts on the number of disciplinary incidents in ninth grade. The third panel of Table 4 reveals that both social value added and test score value added predict fewer incidents, while work hard value added does not. A one standard deviation increase in social value added reduces the number of incidents by 0.0102 compared to only 0.00759 for test score value added, indicating that the SED value added is somewhat more predictive. The ratio of the explained variance using all three value-added measures to the variance explained using test score value added alone is 1.91.

[9] We report a similar pattern of effects on ninth grade GPA in Table S9.

Table 4—Effects on Other Outcomes and Behaviors

	(1)	(2)	(3)	(4)	(5)
On track in ninth grade (114,512 observations)					
Social value added	0.0195			0.0104	0.00911
	(0.00646)			(0.00676)	(0.00675)
Work hard value added		0.0210		0.0130	0.0116
		(0.00765)		(0.00886)	(0.00871)
Test scores value added			0.0141		0.0103
			(0.00421)		(0.00413)
p-value from F-test of work hard and social				0.0125	0.0257
Standard deviation of predicted school effect	0.0193	0.0210	0.0157	0.0217	0.0254
Absences in ninth grade (157,630 observations)					
Social value added	−1.309			−1.108	−1.027
	(0.278)			(0.321)	(0.316)
Work hard value added		−1.062		−0.271	−0.189
		(0.279)		(0.317)	(0.311)
Test scores value added			−0.772		−0.423
			(0.234)		(0.230)
p-value from F-test of work hard and social				0.0000	0.0002
Standard deviation of predicted school effect	1.298	1.060	0.862	1.3088	1.3849
Number of disciplinary incidents in ninth grade (160,148 observations)					
Social value added	−0.0102			−0.0110	−0.0100
	(0.00418)			(0.00518)	(0.00515)
Work hard value added		−0.00677		0.00112	0.00216
		(0.00405)		(0.00468)	(0.00461)
Test scores value added			−0.00759		−0.00536
			(0.00335)		(0.00320)
p-value from F-test of work hard and social				0.0488	0.1106
Standard deviation of predicted school effect	0.0101	0.00676	0.00846	0.0102	0.0117
Ever received school-based arrest (55,564 observations)					
Social value added	−0.00728			−0.00398	−0.00356
	(0.00242)			(0.00317)	(0.00321)
Work hard value added		−0.00766		−0.00480	−0.00382
		(0.00233)		(0.00309)	(0.00298)
Test scores value added			−0.00523		−0.00335
			(0.00180)		(0.00196)
p-value from F-test of work hard and social				0.0032	0.0305
Standard deviation of predicted school effect	0.00722	0.00765	0.00583	0.0081	0.0088

Notes: Results are based on regression of behaviors on out-of-sample value added of social, work hard, and test scores value added. All models include individual demographic controls (race/ethnicity, free and reduced-price lunch, and gender), eighth grade lags (math and English language arts test scores, survey measures, absences, and discipline), and school-level averages for all the demographics and lagged measures as well as year fixed effects. We also include the SES of the student census block proxied by average occupation status and education levels. For each model, we estimate the standard deviation of the predicted school impacts based on the value-added measures included in the model (i.e., $\hat{\sigma}_{\hat{F}}$). In columns 4 and 5, we report the p-value associated with the test that the coefficients on work hard value added and social value added are jointly equal to zero. Missing eighth grade measures were imputed using seventh grade measures and demographic characteristics. Robust standard errors adjusted for clustering at the school level in parentheses.

In sum, schools that raise test scores are not always those that improve SED and vice versa. Schools' SED value addeds are stronger predictors of impacts on all outcomes (excluding test scores) than test score value added. Among the SED dimensions, some schools improve social well-being, while others promote working hard. Consistent with them impacting two different socioemotional dimensions, work hard value added is a stronger predictor of the academic success (as measured by on-track), while social value added is more predictive of impacts on nonacademic behaviors (such as attendance and disciplinary incidents).

V. Impacts on Arrests

Our final key outcome is school-based arrests (bottom panel of Table 4). In models with each value added individually, increasing social value added reduces the likelihood of an arrest by 0.728 percentage points (p-value < 0.01), increasing work hard value added reduces the likelihood of an arrest by 0.766 percentage points (p-value < 0.01), and increasing test score value added reduces the likelihood of an arrest by 0.523 percentage points (p-value < 0.01). The SED value-added measures have greater predictive power than test score value added. The explained variance of school effects on arrests using all three value addeds is 127 percent greater than that using test score value added alone. Because school-based arrests likely have important longer-run implications beyond academic outcomes, these results underscore the importance of evaluating school effects on nonacademic outcomes (Beuermann et al. 2018). Using all three value addeds, the standard deviation of the predicted school effect is 0.88 percentage points. Compared to the average, this represents a relative risk decrease of about 21 percent.

VI. Discussion and Conclusions

We identify persistent school impacts on two distinct dimensions, promoting social well-being and hard work, indicating that surveys can be used to reliably measure SED (net reporting biases) and that schools can foster SED beyond elementary school. This is the first paper to validate school impacts on SED out of sample, provide evidence that the estimates are causal, document potentially instrumental mediating outcomes (e.g., attendance, on track), and show that both crime and longer-run educational attainment outcomes are influenced by attending a high school that improves SED. Moving beyond correlational evidence, this is the first paper to link schools' *causal* impacts on self-reported SED to longer-run outcomes. Our finding that school impacts on SED have larger effects on short- and long-run outcomes than schools' test score impacts has important implications for how policymakers measure school quality.

Our results are consistent with school effects being multidimensional but also with test score and SED value added each being noisy measures of school quality. While this distinction matters conceptually, what matters for policy is that SED value added predicts impacts on longer-run outcomes that are unrelated to test score value added. Relatedly, regarding the usefulness of surveys for policy, note that surveys may be gamed when stakes are attached. As such, while we show that surveys can, and do, measure important educational output in a low-stakes context, our results do not imply that these measures should be used for accountability. Instead, our findings indicate that to better inform policy, further work is needed to (i) derive SED measures that are difficult to manipulate and (ii) identify school practices that improve SED that can be promoted (through training or incentives). While much work remains to be done, our analysis represents an important early step toward understanding how schools may influence SED.

REFERENCES

Alan, Sule, Teodora Boneva, and Seda Ertac. 2019. "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." *The Quarterly Journal of Economics* 134 (3): 1121–62.

Allensworth, Elaine, and John Q. Easton. 2005. *The On-Track Indicator as a Predictor of High School Graduation*. Chicago: University of Chicago Consortium on School Research.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics* 132 (2): 871–919.

Beuermann, Diether W., C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. 2018. "What Is a Good School, and Can Parents Tell? The Multidimensionality of School Output and Parental Preferences." National Bureau of Economic Research Working Paper 25342.

Blackwell, Lisa S., Kali H. Trzesniewski, and Carol Sorich Dweck. 2007. "Implicit Theories of Intelligence Predict Achievement across an Adolescent Transition: A Longitudinal Study and an Intervention." *Child Development* 78 (1): 246–63.

Borghans, Lex, Bas ter Weel, and Bruce A. Weinberg. 2008. "Interpersonal Styles and Labor Market Outcomes." *Journal of Human Resources* 43 (4): 815–58.

Collaborative for Academic, Social, and Emotional Learnings. 2020. "CSI Resources: Frameworks & Competencies." https://casel.org/csi-resources-frameworks-competencies/.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

Credé, Marcus, Michael C. Tynan, and Peter D. Harms. 2017. "Much Ado about Grit: A Meta-analytic Synthesis of the Grit Literature." *Journal of Personality and Social Psychology* 113 (3): 492–511.

Dee, Thomas S., and Emily Penner. 2019. "My Brother's Keeper? The Impact of Targeted Educational Supports." Center for Education Policy Analysis Working Paper 19-07.

Deming, David J. 2014. "Using School Choice Lotteries to Test Measures of School Effectiveness." *American Economic Review* 104 (5): 406–11.

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *The Quarterly Journal of Economics* 132 (4): 1593–1640.

Duckworth, Angela L., Christopher Peterson, Michael D. Matthews, and Dennis R. Kelly. 2007. "Grit: Perseverance and Passion for Long-Term Goals." *Journal of Personality and Social Psychology* 92 (6): 1087–1101.

Dweck, Carol S. 2006. *Mindset: The New Psychology of Success*. New York: Random House.

Dweck, Carol S., and David S. Yeager. 2019. "Mindsets: A View from Two Eras." *Perspectives on Psychological Science* 14 (3): 481–96.

Farrington, Camille A., Melissa Roderick, Elaine Allensworth, Jenny Nagaoka, Tasha Seneca Keyes, David W. Johnson, and Nicole O. Beechum. 2012. *Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance: A Critical Literature Review*. Chicago: University of Chicago Consortium on Chicago School Research.

Fricke, Hans, Susanna Loeb, Robert H. Meyer, Andrew Rice, Libby Pier, and Heather J. Hough. 2019. "Measuring School Contributions to Growth in Social-Emotional Learning." https://edpolicyinca.org/publications/stability-school-contributions-student-social-emotional-learning-gains.

Heckman, James J., and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2): 145–49.

Jackson, C. Kirabo. 2013. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." *Review of Economics and Statistics* 95 (4): 1096–1116.

Jackson, C. Kirabo. 2014. "Teacher Quality at the High School Level: The Importance of Accounting for Tracks." *Journal of Labor Economics* 32 (4): 645–84.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-test Score Outcomes." *Journal of Political Economy* 126 (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: Dataset." *American Economic Review: Insights*. https://doi.org/10.1257/aeri.20200029.

Kaba, Mariame, and Frank Edwards. 2012. *Policing Chicago Public Schools: A Gateway to the School-to-Prison Pipeline*. San Antonio: Project NIA.

Kautz, Tim, James J. Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. 2014. "Fostering and Measuring Skills: Improving Cognitive and Noncognitive Skills to Promote Lifetime Success." National Bureau of Economic Research Working Paper 20749.

Lindqvist, Erik, and Roine Vestman. 2011. "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment." *American Economic Journal: Applied Economics* 3 (1): 101–28.

Loeb, Susanna, Michael S. Christian, Heather J. Hough, Robert H. Meyer, Andrew B. Rice, and Martin R. West. 2018. "School Effects on Social-Emotional Learning: Findings from the First Large-Scale Panel Survey of Students." https://files.eric.ed.gov/fulltext/ED591089.pdf.

National Student Clearinghouse. 2020. StudentTracker for High Schools. National Student Clearinghouse.

Revelle, W. 2007. "Experimental Approaches to the Study of Personality." In *Handbook of Research Methods in Personality Psychology*, edited by R. W. Robins, R. C. Fraley, and R. F. Krueger, 37–61. New York: Guilford Press.

Rimfeld, Kaili, Yulia Kovas, Philip S. Dale, and Robert Plomin. 2016. "True Grit and Genetics: Predicting Academic Achievement from Personality." *Journal of Personality and Social Psychology* 111 (5): 780–89.

Sisk, Victoria F., Alexander P. Burgoyne, Jingze Sun, Jennifer L. Butler, and Brooke N. Macnamara. 2018. "To What Extent and under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-analyses." *Psychological Science* 29 (4): 549–71.

University of Chicago Consortium on School Research. 2020. Archive of Data on Chicago Public Schools 2011–2017. University of Chicago Consortium on School Research.

Waddell, Glen R. 2006. "Labor-Market Consequences of Poor Attitude and Low Self-Esteem in Youth." *Economic Inquiry* 44 (1): 69–97.

Walton, Gregory M., and Shannon T. Brady. 2017. "The Many Questions of Belonging." In *Handbook of Competence and Motivation: Theory and Application*, edited by A. J. Elliot, C. S. Dweck, and D. S. Yeager, 272–93. New York: Guilford Press.

West, Martin R., Matthew A. Kraft, Amy S. Finn, Rebecca E. Martin, Angela L. Duckworth, Christopher F. O. Gabrieli, and John D. E. Gabrieli. 2016. "Promise and Paradox: Measuring Students' Noncognitive Skills and the Impact of Schooling." *Educational Evaluation and Policy Analysis* 38 (1): 148–70.