

SCHOOL DISTRICT REFORM IN NEWARK: WITHIN- AND BETWEEN-SCHOOL CHANGES IN ACHIEVEMENT GROWTH

MARK CHIN, THOMAS J. KANE, WHITNEY KOZAKOWSKI,
BETH E. SCHUELER, AND DOUGLAS O. STAIGER*

In the 2011–12 school year, the Newark Public School district (NPS) launched a set of educational reforms supported by a gift from Facebook CEO Mark Zuckerberg and Priscilla Chan. Using data from 2008–09 through 2015–16, the authors evaluate the change in Newark students' achievement growth relative to similar students and schools elsewhere in New Jersey. They measure achievement growth using a “value-added” model, controlling for prior achievement, demographics, and peer characteristics. By the fifth year of reform, Newark saw statistically significant gains in English language arts (ELA) achievement growth and no significant change in math achievement growth. Perhaps because of the disruptive nature of the reforms, growth declined initially before rebounding in later years. Much of the improvement was attributed to shifting enrollment from lower- to higher-growth district and charter schools.

In the fall of 2010, Facebook CEO Mark Zuckerberg, New Jersey Governor Chris Christie, and Newark Mayor Corey Booker announced a school improvement effort in the Newark Public School district (NPS), to be aided by a \$100 million gift from Zuckerberg and Priscilla Chan. This

*MARK CHIN is a PhD Candidate at Harvard University. THOMAS J. KANE is the Walter H. Gale Professor of Education at Harvard University. WHITNEY KOZAKOWSKI is a PhD Candidate at Harvard University. BETH E. SCHUELER is a Postdoctoral Research Fellow at Harvard University. DOUGLAS O. STAIGER is the John Sloan Dickey Third Century Professor at Dartmouth College.

This project was funded by a grant from the Startup:Education Foundation (now the Chan-Zuckerberg Initiative). We thank Bari Erlichson, Kendra Carter, Kate Fletcher, Jesse Margolis, Sarah McLean, James Riddlesperger, and Peter Shulman for their help explaining the context in Newark and accessing the New Jersey data. Kirsten Slungaard Mumma helped analyze the data, and Jon Fullerton managed the project. Although unrelated to the current project, we disclose the following interests: Kane served as an expert witness for Gibson, Dunn, and Crutcher LLP to testify in *Vergara v. California*. Staiger is cofounder, has an equity interest in, and consults for ArborMetrix, a company that sells performance measurement systems and consulting services to health care providers. For information regarding the data and/or computer programs utilized for this study, please address correspondence to the authors at doug.staiger@dartmouth.edu.

KEYWORDS: school district reform, student performance, productivity, value-added models, cross-sectional time-series analysis

gift—combined with \$100 million in matched funds from other donors—represented approximately 4% of Newark’s school budget over the five years of the grant (2011–12 to 2015–16). Yet, it provided the city and the district with the flexibility to implement an ambitious slate of reforms. With the appointment of a new school superintendent, Cami Anderson, the reforms were launched in the fall of the 2011–12 school year. We provide a summary of the results through the reforms’ fifth year.

“Reform” in Newark was not a single intervention but rather a package of measures that included a new teacher contract, new school leaders, a new curriculum aligned to the Common Core standards, school turnaround efforts (known as “renew schools”), charter school expansion, school closures, and, eventually, a universal choice plan (which allowed parents to submit a single application for their child(ren) to attend a district or local charter school). Rather than attempt to parse the effect of each reform component (likely an impossibility), we divide the reforms into two broad categories: “within-school” reforms (such as personnel changes, Common Core implementation, and school turnaround efforts aimed at improving the results of existing schools) and “between-school” reforms (such as school closures, charter expansion, and universal choice, aimed at re-allocating students toward more effective schools). Using a value-added model to compare the achievement growth of students in Newark to similar students attending similar schools elsewhere in New Jersey, we decompose the change in average annual achievement growth into its within-school and between-school components.

A similar framework has been used to measure the sources of productivity growth in other countries (e.g., Bartelsman, Haltiwanger, and Scarpetta 2013) and industries (e.g., Chandra, Finkelstein, Sacarny, and Syverson 2016). In those studies, productivity growth was often driven by gains in market share by more efficient firms—as opposed to productivity improvements within existing firms. In K–12 education, the contributions from changing market share are typically blunted when school assignment is based on the location of a student’s residence. By closing some schools, allowing students to move to charters, and instituting a universal choice plan, however, the Newark reform strategy allowed parents to switch schools without switching residences. We assess the degree to which this opportunity improved the system’s overall productivity.

Compared to the two academic years preceding the reforms (those ending in spring 2010 and spring 2011), we find that by spring 2016 achievement growth in Newark improved significantly in English language arts (ELA) and showed no significant change in mathematics. Overall progress was not linear. In the initial years of the reform, annual average achievement growth declined in math and ELA, in both the charter and the district schools, only to increase in later years. After bouncing back in 2014–15, achievement growth in the average Newark school is now significantly

higher than for similar students in the rest of the state in ELA and higher, though not significantly so, in math.

A distinguishing characteristic of the Newark reforms is the shift in enrollment toward higher value-added schools. The relationship between achievement growth and shifting enrollment is much stronger in Newark than elsewhere in New Jersey. The shifts occurred for several reasons, including the closure of schools with below-average achievement growth, expansion of charter schools with higher growth, and the introduction of universal choice (allowing parents to apply for district schools, magnet schools, and charter schools with a single application). Indeed, 61% of Newark's improvement in English achievement growth between the baseline years and 2015–16 was attributable to shifting enrollment. In mathematics, the sole source of improvement was between-school movement, as math achievement growth would have declined in Newark relative to the rest of the state if not for enrollment shifts toward schools with faster achievement growth in math.

Our article adds to a growing literature on school and district turnarounds (e.g., Gill, Zimmer, Christman, and Blanc 2007; Young et al. 2009; Dee 2012; Fryer 2014; Papay 2015; Abdulkadiroglu, Angrist, Hull, and Pathak 2016; Harris and Larsen 2016; Heissel and Ladd 2016; Strunk et al. 2016; Schueler, Goodman, and Deming 2017; Zimmer, Henry, and Kho 2017). As other research has begun to show, turnaround efforts can produce both positive and negative externalities for student achievement growth. For instance, school choice and closures may increase pressure on schools to improve. Conversely, large movements of students and teachers can be disruptive, both for students who move and for their new classmates. By examining district-wide productivity change, our study incorporates both types of spillover effects.

Furthermore, existing studies tend to focus on the impact of reform on students in schools while they are undergoing turnaround. Less attention has been paid to understanding how disruptions in one set of schools could affect other schools and eventually alter system-wide productivity for future cohorts. For example, several studies examined the effect of school closure on displaced students (e.g., de la Torre and Gwynne 2009; Engberg, Gill, Zamarro, and Zimmer 2012; Brummet 2014; Kemple 2015; Carlson and Lavertu 2016), but few considered the impact of closure on students who avoid attending a low-performing school because it had been closed prior to their entry into the district. By contrast, we apply a framework that examines the relative contributions of various turnaround mechanisms—within-school improvements and between-school reallocation of students—to overall district-wide productivity change over time.

Newark Reforms

The Newark Public School district (NPS) has long been the focus of reform efforts. In its 1985 *Abbott v. Burke* ruling, the New Jersey State Supreme

Court declared the state's reliance on local property taxes for financing K–12 schools unconstitutional, and classified Newark as one of 28 districts needing additional state assistance (Hess 2007).¹ These “Abbott districts” are characterized by poor academic performance and high concentrations of economically disadvantaged students. NPS was placed under state control in 1995 after the New Jersey Department of Education determined Newark was not meeting district certification standards (Morel 2018). It was not until September 2017 that the state Board of Education voted to fully restore local control. Therefore, the district was under state takeover for the full period we study.

In 2010, then Mayor Cory Booker and Governor Chris Christie began collaborating on an effort to reform the Newark school system. They found a willing investor in Mark Zuckerberg, chief executive of the social networking website Facebook. Zuckerberg and Priscilla Chan (cofounder of the Chan-Zuckerberg Initiative) committed \$100 million to NPS, matched by an additional \$100 million from primarily foundations and private donors (Fulbeck et al. 2016). Zuckerberg, Booker, and Christie announced the donation on the Oprah Winfrey show in September 2010. To put the magnitude of the donation in perspective, the \$200 million (which was to be spent over five years) was approximately 4% of the district's five-year \$5 billion budget. To carry out the reforms, Christie appointed Cami Anderson as superintendent of schools in May 2011. Anderson had led New York City's alternative high school program under Chancellor Joel Klein. We provide a timeline of the Newark reforms in Table 1.

The reforms can be grouped into two broad categories: those that aimed to improve existing schools (within-school reforms) and those that aimed to reallocate students to more effective schools (between-school reforms).

Within-School Reforms

Over her first three years as superintendent, Anderson replaced more than half of the district's principals, including one-third in her first year (Russakoff 2015). She reorganized the central office, cutting 120 positions and bringing in new leadership (Meyer 2013). The new team implemented district-wide reforms such as monthly training sessions with principals, an online platform to facilitate mutual consent hiring, an early warning system to prevent drop out, changes to data and accountability systems (e.g., school data dashboards), a revamped student registration system, and extended learning time in a subset of schools (Russakoff 2015). Newark also piloted blended learning models, incorporating educational software in classrooms (NPS 2013).

In November 2012, the district ratified a new contract with the Newark Teachers Union (NTU). The four major contract provisions included 1) a

¹There are now 31 such districts.

Table 1. Timeline of Recent Events and School Reforms in Newark

1985	First <i>Abbott v. Burke</i> State Supreme Court finance ruling leads to identification of 31 disadvantaged “Abbott districts”
1995	State takes control of Newark schools
Sept. 2010	\$100 million StartUp:Education donation announced on Oprah Winfrey Show
May 2011	Governor Chris Christie appoints Cami Anderson as Newark superintendent of schools
Spring 2012	Decision to consolidate twelve schools into eight “renew” schools and expand pre-K announced
2011–12	Nearly 25% of principals replaced leading up to the 2011–12 school year Central office reorganized New district-wide systems for registration, hiring, and data and accountability implemented Blended learning models and interim assessments with data use training piloted Extended learning time implemented in subset of schools
Spring 2012	Four K–8 schools closed at end of school year
Sept. 2012	Eight K–8 schools “renewed” (four absorb students from closed schools)
Nov. 2012	Agreement reached on new pay-for-performance teacher contract
Sept. 2013	A subset of high schools “renewed”
Spring 2013	Three K–8 schools closed at end of school year
Fall 2013	“One Newark” universal enrollment plan announced
Spring 2014	Families submit school choice preferences for 2014–15 placements Seven K–8 schools closed at end of school year
Sept. 2014	Eight additional K–8 schools “renewed”
Spring 2015	Families submit “Newark Enrolls” school choice preferences for 2015–16 placements
Spring 2015	First administration of Common Core–aligned PARCC exam in New Jersey
Summer 2015	Cami Anderson resigns and Chris Cerf assumes superintendent role
Summer 2016	State announces plan to fully restore local control of Newark schools by 2017–18

new system to evaluate teachers (combining classroom observations, student growth, and other indicators such as lesson plans and teacher attendance); 2) differentiated teacher compensation, including incentives for high-performing teachers to stay in low-performing schools; 3) extended learning time in 28 schools; and 4) greater school-based decision making (including the ability for teachers to vote to overturn portions of the collective bargaining agreement). In return, NTU members received a \$31 million one-time payment to resolve outstanding wage demands from prior years and \$20 million in stipends during the first year of implementation. A majority of teachers perceived the new evaluation system to be fair and accurate, and reported that the extended learning time was helpful to students (Fulbeck et al. 2016). Although teachers had mixed support for differentiated compensation, Newark was able to retain teachers with higher ratings at higher rates under the new contract (Fulbeck et al. 2016).

Newark also began emphasizing the Common Core state standards before many other New Jersey districts, adopting Common Core–aligned math and literacy instructional materials in grades K–8 starting in 2013–14. For example, nearly 90% of NPS schools serving grades 3–8 adopted the Expeditionary Learning (EL) curriculum in English. These resources have been highly rated by EdReports.org for Common Core alignment. The

change in curriculum may have contributed to temporary declines in performance on the New Jersey Assessment of Skills and Knowledge (NJASK) test, as the materials were less aligned with NJASK and teachers and students likely needed time to learn how to use these new resources effectively.

The final major within-school reform was the attempted turnaround of underperforming schools. Eight K–8 schools were chosen for turnaround during the 2012–13 academic year, and an additional eight were chosen for 2014–15.² The schools—labeled “renew schools”—were the focus of a multi-pronged effort. First, principals were required to reapply for their jobs; in the first round, half of the principals were replaced (Russakoff 2015). Once hired, principals were able to rehire or replace teachers (Calefati 2012; McGlone 2013). About half of the teachers were replaced in the first round.³ Beyond staffing changes, students were given extended learning time and teachers received additional professional development opportunities. Renew schools also conducted extra outreach to families and offered students better access to nurses, social workers, and community-based mentoring (Kamenetz 2013; Fulbeck et al. 2016).

Between-School Reforms

While working to improve the district schools, Newark leaders also sought to shift students into higher-quality schools. School closures were one component of that strategy. Between 2011–12 and 2014–15, 11 traditional district schools serving grades K–8 were closed based on low enrollment and poor performance.⁴ Some students from closed schools transferred to charter schools; others were reassigned to nearby district schools, some of which were renew schools. Four of the eight schools in the first round of renewal received students from one of the four schools that closed that year. In addition to district school closures, three charter schools were closed between 2011–12 and 2014–15 for several reasons, including poor academic performance and financial and legal concerns (Mooney 2013). Students affected by the 2013–14 school closures were offered transportation from several shuttle hub locations to their new schools. The district also opened seven new schools, including two single-sex schools, to serve students in grades 6–12.

Meanwhile, enrollment in Newark’s charter schools grew rapidly. In spring 2011, 14% of all NPS students were enrolled in charters. By 2015–16, the percentage had doubled to 28% and was even higher among K–8 students (32%). The charter sector included 20 different operators by 2015–16. The largest were TEAM Academy, affiliated with the national

²A small number of high schools were renewed in 2013–14 as well, but our analysis focuses on K–8 renew schools.

³Teachers who were not rehired were provided other employment options within the district.

⁴We define closed schools as those with school codes no longer in use and with facilities no longer used by the school that previously resided there.

Knowledge is Power Program (KIPP) network, and North Star Academy, affiliated with the national Uncommon Schools network. These two operators accounted for 46% of Newark's charter enrollment in 2015–16 and 56% of the growth in charter enrollment between 2010–11 and 2015–16.

In the winter and spring of 2014, Newark instituted a universal choice system (now known as “Newark Enrolls”). This system allowed families to rank their preferred schools on a single application, choosing from traditional district schools, magnet schools, and nearly all of Newark's charter schools. Students were then matched to schools.⁵ Previously, families were assigned to a district school based on the location of their residence; if they wanted their children to attend a charter school or a magnet school, they were required to submit a separate application to each school. Beyond the shuttles mentioned above, NPS directly provided transportation only to students with an Individualized Education Program (IEP) and provided public bus tickets to students living two or more miles from their school (NPS 2016). TEAM Academy (KIPP) was the only charter network that provided busing (Russakoff 2015).

Despite the national attention on Newark, the impact of the reforms on student achievement growth has not previously been evaluated.⁶ We focus on annual achievement growth—rather than levels of achievement—given the changes in the composition of students attending district and charter schools in Newark.

Data

Our data cover all New Jersey public school students in grades 4–8 for seven academic years (2009–10 through 2015–16). The data include student demographic and program participation indicators including age, gender, race or ethnicity, limited English proficiency status (LEP), eligibility for free or reduced-price lunch (FRPL), and disability classification, as well as current- and prior-year test performance on state mathematics and English language arts (ELA) standardized tests.

For the academic years ending in 2009 through 2014, our student outcomes are NJASK state test scores.⁷ In spring of 2015 and 2016, New Jersey

⁵In addition to accounting for families' rankings, the algorithm prioritized keeping siblings together, allowing students to attend a school in their residential neighborhood, and increasing the representation of students with an Individualized Education Program (IEP) or free or reduced-price lunch (FRPL) eligibility in schools where these students are underrepresented in the applicant pool relative to the citywide average. For magnet schools, the matching also accounts for school rankings of applicants.

⁶Baker and Weber (2015) used publicly available school-level data from grades 6–8 to compare trends in average scale scores (not growth) from 2009 to 2014 in Newark to the rest of New Jersey, and found no difference.

⁷The distribution of NJASK scale scores across grades and years revealed slight ceiling effects for math. As such, we calculated rank-based standardized scores (also known as van der Waerden scores [Conover 1999]) for math, standardized by grade and year. Because we observed no apparent ceiling effect in the NJASK ELA scale scores or in the PARCC math and ELA scores, we standardized by grade and year using the usual method (subtracting the mean and dividing by the standard deviation).

Table 2. Student Sample Characteristics, 2009–10 to 2015–16

<i>Characteristic</i>	<i>Newark district</i>	<i>Newark charter</i>	<i>Other Abbotts</i>	<i>Rest of New Jersey</i>
<i>N of students</i>	11,716	3,698	80,030	365,680
<i>N of schools</i>	52	15	328	1,421
<i>Male</i>	0.512	0.460	0.508	0.512
<i>White</i>	0.079	0.007	0.107	0.626
<i>African American</i>	0.493	0.797	0.325	0.103
<i>Asian</i>	0.009	0.009	0.034	0.112
<i>Latino</i>	0.417	0.164	0.498	0.134
<i>American Indian</i>	0.001	0.000	0.001	0.001
<i>Multiracial/Other</i>	0.001	0.023	0.034	0.024
<i>FRPL eligible</i>	0.883	0.829	0.798	0.235
<i>LEP</i>	0.053	0.004	0.072	0.012
<i>Has disability</i>	0.165	0.106	0.147	0.156
<i>Prior math achievement</i>	-0.618	-0.166	-0.482	0.157
<i>Prior ELA achievement</i>	-0.668	-0.158	-0.550	0.167

Notes: Sample sizes are averaged across all years for students with valid ELA scores. The numbers with valid math scores were similar. Charter networks count as a single school in these data. FRPL, free or reduced-price lunch; LEP, limited English proficiency; ELA, English language arts.

administered the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment, designed to reflect the Common Core State Standards. For both tests, we standardized scores by grade and year. We limited our analysis to students with non-missing outcomes and covariates. We also dropped those in schools and grades with fewer than five students in a year.

In Table 2, we report mean characteristics for four subgroups: students attending Newark district schools, Newark charter schools, other Abbott district schools, and the remainder of New Jersey schools. Similar to the students in the other Abbott districts, students in Newark district schools were more likely to be African American or Latino and more likely to be eligible for subsidized lunches than were students in the rest of New Jersey. They were also lower achieving, with mean standardized test scores in the prior year two-thirds of a standard deviation below the state average in mathematics and ELA. Compared to students in Newark district schools, students in Newark charters were more likely to be African American, but less likely to be Latino, special education, or LEP students. Newark charter school students had somewhat higher achievement in the prior year than did the Newark district students, scoring 0.16 standard deviations (SD) below the statewide average, rather than two-thirds of a SD below.

Methods

To assess the impact of the Newark reforms, we measure the rate of annual achievement growth for all Newark students (including those attending charter schools) relative to other students in New Jersey with similar

baseline achievement and demographics who attend schools with similar mean baseline achievement and demographics.

Specifically, we estimate the following value-added model:

$$(1) \quad A_{ijt} = \beta_0 + X_{ijt}\beta_1 + \delta_{jt} + \varepsilon_{ijt}$$

where A_{ijt} represents the score on the mathematics or ELA test for student i in district j at time t . The term X_{ijt} represents our vector of covariates, including student demographic characteristics, prior test performance, the average characteristics of peers in the school and grade, grade-by-year fixed effects, and (especially given the change to the new test in 2015) grade-by-year interactions with the student and peer covariates. The term δ_{jt} represents district-by-year fixed effects. Although New Jersey treats charter schools as districts of their own, we categorize charter schools in Newark and the Abbott districts with their neighboring district when estimating Equation (1).

We include district-by-year fixed effects to avoid conflating the influences of student and school-level characteristics (for which we seek to adjust) and the effect of district-level policy changes (which we seek to measure). In other words, we estimate the effect of student and school characteristics based only on within-district-year variation in those covariates, which may be correlated with district-level policy changes. (For instance, other low-income districts may have been pursuing the same policies as Newark.) Accordingly, our estimates adjust Newark's performance relative to other schools based solely on the measured effect of those school characteristics within districts in each year.

We then estimate the difference in adjusted achievement growth between Newark and other New Jersey districts with the following second-step equation:

$$(2) \quad A'_{ijt} = \gamma_0 + \gamma_{2010-2011} \text{Newark}_j + \sum_{t=2012}^{2016} \gamma_t \text{Newark}_j * \text{Year}_t + \tau_t + u_{ijt}$$

where A'_{ijt} is student i 's achievement, adjusted by subtracting the estimated effect of prior achievement, demographics, and mean characteristics ($X_{ijt}\beta_1$) from Equation (1), leaving the district-by-year effects and the residuals ($A'_{ijt} = \delta_{jt} + \varepsilon_{ijt}$). This two-step method is analogous to the approach used in Chetty, Friedman, and Rockoff (2014), in which the authors used only within-teacher variation to estimate the effects of covariates (Equation (1)), and then analyzed the between-teacher variation in a second stage (Equation (2)).

The coefficient in Equation (2) on the Newark indicator, $\gamma_{2010-2011}$, captures the adjusted achievement gain in Newark relative to similar students and schools elsewhere in New Jersey in the pre-reform years (2010 and 2011). The remaining coefficients, $\gamma_{2012} \dots \gamma_{2016}$, represent the differential change in adjusted achievement growth in Newark from the pre-reform years, relative to the change in other New Jersey districts. To capture

changes in achievement growth outside Newark, we include year fixed effects (τ_t). For both equations, we calculate standard errors by clustering at the school level. Additionally, we estimate similar equations including separate time trends for Newark charters, renew, and other district schools.

Within-School Changes in Achievement Growth

We also explore the within-school changes in achievement growth in Newark by estimating a model identical to the one depicted in Equation (2), except that we 1) include school fixed effects, μ_s , and 2) omit the main effect of enrollment in Newark (i.e., the baseline value-added of Newark in Equation (2), $NEWARK_j$) because of collinearity of the Newark indicator with the school fixed effects. Thus, our coefficients of interest— γ_t —represent the average change in productivity within Newark schools over time. We also explore within-school productivity trends of various school types (charters, non-renew traditional district, and renew traditional district schools).

Enrollment Growth and Achievement Growth

Given Newark's efforts to allow children to move between schools, we estimate the relationship between baseline measures of school achievement growth and subsequent school enrollment changes. Specifically, we follow Chandra et al. (2016) and estimate the following model at the school level, by year and subject, separately for Newark and for comparison districts:

$$(3) \quad \Delta Enrollment_{sj,t+1} = \lambda_0 + \lambda_1 \hat{\mu}_{sjt} + \alpha_j + \mu_{sjt}$$

where $\Delta Enrollment_{sj,t+1}$ is the difference between school s 's enrollment in time $t+1$ and the school's enrollment in time t , while $\hat{\mu}_{sjt}$ is an estimate of school s 's value-added in year t , and α_j is a district fixed effect.⁸ The coefficient of interest, λ_1 , captures the within-district relationship between school value-added in time t and the subsequent change in enrollment between time t and $t+1$. A positive value of λ_1 would indicate that higher value-added schools gained market share. We first present results using enrollment in grades from which school value-added is constructed (grades 4–8). Because Newark has many K–8 schools, however, the earliest grades may be most sensitive to increased availability of schooling options. To account for this possibility, we present results using enrollment in grades 1–3.⁹ We

⁸We estimate school value-added in year t first by estimating Equation (1) across all students in New Jersey. Then, we regress A'_{ijt} on school random effects, estimating each model one year at a time. Since we use value-added as a predictor in Equation (3), we estimate empirical Bayes shrunken random effects for schools to account for differences in the reliability of estimates across schools due to enrollment differences (Kane and Staiger 2008).

⁹We exclude kindergarteners because kindergarten enrollments have anomalies indicating data quality issues in some years in the data.

estimate the model using three samples—Newark, other Abbott districts, and the rest of New Jersey—to compare the relationship between school productivity and enrollment growth observed in Newark to other districts in the state.

Finally, we decompose the change in district-wide average school value-added to assess the extent to which productivity changes in Newark were attributable to within-school change versus between-school movement. We follow Chandra et al. (2016) in their adaptation of Foster, Haltiwanger, and Krizan (2001), Foster, Haltiwanger, and Syverson (2008), and Baily, Hulten, and Campbell (1992). Our decomposition takes the following form:

$$\begin{aligned}
 (4) \quad \Delta \bar{q}_t = & \underbrace{\sum_{s \in C_t} \theta_{s,t-1} \Delta q_{s,t}}_{\text{within}} + \underbrace{\sum_{s \in C_t} (q_{s,t-1} - \bar{q}_{t-1}) \Delta \theta_{s,t}}_{\text{between}} + \underbrace{\sum_{s \in C_t} \Delta q_{s,t} \Delta \theta_{s,t}}_{\text{cross}} \\
 & + \underbrace{\sum_{s \in M_t} \theta_{s,t} (q_{s,t} - \bar{q}_{t-1})}_{\text{entry}} + \underbrace{\sum_{s \in X_t} \theta_{s,t-1} (\bar{q}_{t-1} - q_{s,t-1})}_{\text{exit}}.
 \end{aligned}$$

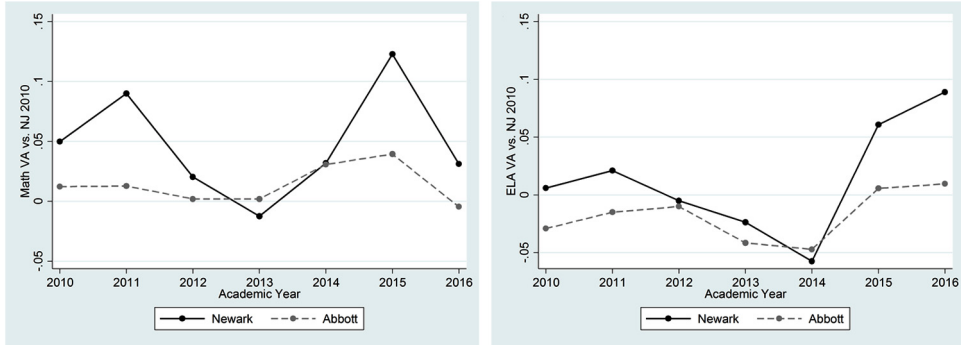
In Equation (4), \bar{q}_t is the district’s average value-added in a given subject in school year t , and Δ represents change.¹⁰ Therefore, the outcome is the difference in average value-added between two periods ($t - 1$ and t). On the right side, $q_{s,t}$ is value-added for school s in year t , and $\theta_{s,t}$ is the share of students in the district enrolled in school s in year t . Term C_t is the set of schools that were open in both $t - 1$ and t , M_t is the set of schools that opened in year t , and X_t are schools that closed between $t - 1$ and t .

The above model decomposes average change in a district’s productivity into five terms. The *within* term reflects changes in average value-added in the district attributable to value-added improvements among schools holding their share of the district’s enrollment constant. The remaining terms reflect the various ways students reallocate across schools. The *between* term represents the extent to which productivity improvements resulted from students moving from lower value-added schools to schools that were already higher value-added in year $t - 1$.

The *cross* term represents enrollment growth at schools with value-added that improved between the two periods. If schools that were growing in enrollment were also seeing declines in the average achievement growth, this term will be negative. We categorize the *cross* term as a “between school” movement, since it captures the degree to which enrollment shifted toward improving schools. (Some authors categorize this term as neither *within* nor *between*, but as its own category.)

¹⁰To recover district and school average value-added within a year for the decomposition analysis, we first estimate Equations (1) and (2) for both subjects across all New Jersey students. Then, we take A'_{ijt} from Equation (2) and subtract the year effects (τ_t) estimated in Equation (2) to determine student-level value-added (to remove remaining year effects in the district-year fixed effects estimated in Equation (1)). We average this result to the school and district level for each year.

Figure 1. Trends in Math and ELA Value-Added in Newark and the Abbott Districts Relative to the Rest of New Jersey



Notes: Value-added calculated by first estimating Equation (1), taking the residuals from that regression and adding to them the estimated district-year fixed effects (i.e., A'_{ijt} in the text description). Then, in an equation similar to Equation (2), we regress A'_{ijt} on the interactions between a Newark indicator and year dummies, interactions between an Abbott indicator and year dummies, and year fixed effects. We plot coefficients from this second regression here. ELA, English language arts; VA, value-added.

The *entry* term captures improvements in value-added due to new schools opening with higher value-added than the previous district average. The *exit* term captures productivity gains attributable to lower than average value-added schools closing. Each of these terms could be negative if changes in value-added, enrollment shares, or the district's portfolio of schools contributed to reductions in average district-wide value-added.

We decompose the change in value-added into a “long difference” in which \bar{q}_t represents the change in average value-added between the two baseline years (2010 and 2011) and our final year (2016) in a given district. We also decompose the change in value-added between each consecutive year in our time series to explore whether the *within* and *between* contributions varied over time. We calculate these terms for Newark and the other Abbott districts, treating the Abbotts as a single district for the decomposition. As a result, the *between* term for the Abbott districts captures student movement across both schools and districts.

Results

Figure 1 portrays the trend in district-by-year effects ($\widehat{\delta}_{j,t}$) for Newark and the average of the other Abbott districts relative to the rest of the state. The first year, ending in spring 2010, occurred before the \$100 million donation was announced in September 2010. The second year occurred after the gift was announced but before the appointment of the new superintendent and before major reforms began in Newark. Thus, we consider 2010 and 2011 as the two baseline years, and 2012 through 2016 as the five years post-reform.

Relative to the rest of the state and the Abbott districts, Figure 1 illustrates that value-added in Newark declined (or, at the very least, showed no

obvious improvement) in the first years of the reform (2012–2014) in both math and ELA. However, value-added increased sharply for Newark in 2015 in both subjects. In that single year, Newark's achievement growth increased by 0.09 SD in math (from 0.03 to 0.12) and 0.11 SD in ELA (from –0.05 to 0.06). Compared to the largest districts in New Jersey over this time, Newark's change was larger than 1 SD of the change that the average large district experienced.¹¹

Newark was not the only low-income district to see achievement growth increase with the administration of the new Common Core–aligned PARCC assessments in 2015. During the same year, growth in the Abbott districts also increased (although less than in Newark) by 0.01 and 0.06 SD in mathematics and English, respectively. In data published by the state, both Newark and the average Abbott district also saw large increases in their 2015 Student Growth Percentiles in English and math.¹² In 2016, achievement growth in Newark remained well above the state and the Abbott districts in English, but declined in math (remaining above the Abbott districts and the state, but no longer significantly so).

Table 3 reports estimates from Equation (2) showing the achievement growth trends, overall and separately for traditional and charter schools in Newark.¹³ The first row of the table reports the difference between Newark and the rest of the state in value-added in the two years prior to the reforms. As reported in column (1), the combined achievement growth of Newark district and charter schools in math was 0.068 SD higher than for similar students elsewhere in New Jersey pre-reform. In column (4), the 0.017 SD difference in ELA growth was not statistically different from zero. As reported in columns (3) and (6), however, the Newark advantage in achievement growth before the reforms was primarily attributable to the Newark charter schools, for which annual achievement growth was 0.319 SD higher in math and 0.215 SD higher in ELA. To put the magnitude of these estimates in context, one standard deviation in value-added across the largest statewide districts was .08 in math and .05 in ELA in the baseline years.

¹¹We define the largest districts as those that enroll at least 3,000 tested students in grades 4–8. This approach yields a total of 31 districts.

¹²A student growth percentile (SGP) in year t is the percentage of students at the same grade level statewide with similar performance on the $t - 1$ test that the student outperformed in year t . Therefore, a median school SGP greater than 50 indicates that the median student outperformed more than half of her peers statewide with the same baseline score the previous spring. The weighted average of median SGPs in ELA in the Newark district schools jumped from 40 in 2014 to 49 in 2015 to 54 in 2016. It could be that a portion of the jump in SGP scores in Newark—especially in ELA—was an artifact of a transition to the PARCC. Other districts serving low-income students also witnessed a jump in SGPs in 2015. While we adjust for the mean characteristics of a student's schoolmates (including the percentage receiving FRPL) using within-district variation and allow those adjustments to vary by year when estimating value-added, SGP models adjust only for a student's baseline achievement.

¹³We estimate the effect of district and charter schools in one pooled model for math (columns (2)–(3)) and one pooled model for ELA (columns (5)–(6)). The model includes a main effect for Newark district schools and Newark charter schools, Newark district school-by-year indicators, Newark charter school-by-year indicators, and year fixed effects.

Table 3. Trends in Value-Added in Newark Schools Relative to the Rest of New Jersey

Variable	Math			ELA		
	All (1)	District (2)	Charter (3)	All (4)	District (5)	Charter (6)
Newark vs NJ difference in 2010/2011	0.068*** (0.023)	0.028 (0.018)	0.319*** (0.033)	0.017 (0.018)	-0.014 (0.014)	0.215*** (0.035)
Change in Newark vs NJ difference relative to 2010/2011						
2012	-0.048** (0.020)	-0.050** (0.021)	-0.136*** (0.046)	-0.021 (0.018)	-0.022 (0.019)	-0.090** (0.045)
2013	-0.081*** (0.022)	-0.092*** (0.024)	-0.171*** (0.038)	-0.034** (0.015)	-0.046*** (0.016)	-0.097*** (0.027)
2014	-0.042 (0.026)	-0.061*** (0.022)	-0.126** (0.063)	-0.066*** (0.018)	-0.082*** (0.017)	-0.133*** (0.035)
2015	0.047 (0.035)	0.035 (0.027)	-0.095 (0.084)	0.043* (0.024)	0.031* (0.018)	-0.066 (0.057)
2016	-0.036 (0.025)	-0.057** (0.024)	-0.172*** (0.040)	0.070*** (0.026)	0.058*** (0.021)	-0.047 (0.048)

Notes: Robust standard errors in parentheses (clustered at the school level). Estimates generated by a two-step process in which we first estimated a value-added model controlling for student demographics (indicators for race/ethnicity, gender, special education status, limited English proficiency status, and free and reduced price lunch status), prior achievement (including cubic polynomials of math and ELA achievement scores interacted with grade), peer covariates (school-grade cohort means for all demographic and prior achievement characteristics), district-by-year fixed effects, grade-by-year fixed effects, and interactions between our grade-by-year fixed effects and all student and peer covariates. We then estimate a second equation in which the outcome is the residual plus district-by-year fixed effects estimated in the first equation, and the predictors are a main effect for Newark, Newark-by-year indicators, and year fixed effects. For columns (2)–(3) and (5)–(6), in the second equation we include a main effect for Newark district schools and Newark charter schools, Newark district school-by-year indicators, Newark charter school-by-year indicators, and year fixed effects. ELA, English language arts. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

In other words, Newark charters were more than one SD above the highest district value-added in math and almost two SDs above the highest district in ELA. By contrast, the Newark district schools had comparable achievement growth to similar schools elsewhere in New Jersey in both subjects in the two baseline years.

The Newark charter schools' advantage in achievement growth on the eve of the Newark reforms was quite large—two to three times as large as the difference in achievement growth associated with having a novice teacher versus an experienced teacher. As other research suggests, Newark is home to one of the most effective charter sectors in the nation in terms of student growth on standardized exams (CREDO 2015). The only other city known to have a similarly effective charter sector is Boston, Massachusetts (Angrist et al. 2010, 2012; Abdulkadiroglu et al. 2011). As we discuss below, the close proximity of an unusually effective charter sector played a major role in the impact of the Newark reforms.

The remaining rows of Table 3 report the change in Newark's value-added during the reform years relative to the rest of New Jersey. In the first three years of reform, Newark's achievement growth declined in mathematics, in both district and charter schools. By 2013, achievement growth in math was 0.092 SD lower in district schools and 0.171 SD lower in charter schools than in the pre-reform years.

By 2015, across all Newark schools, achievement growth recovered to be a bit above pre-reform levels, although the difference was only marginally significant for ELA. Achievement growth by 2015 was slightly higher than in the baseline years for district schools and slightly lower for charter schools.

The improvements in ELA achievement growth persisted and were significantly above baseline by 2016, although some backsliding occurred in math, with achievement growth significantly below the baseline years in both the charter and the district sectors. Yet, despite the decline within both sectors, when the charter and district schools are combined in column (1), the Newark achievement growth advantage in math in 2016 was not statistically different from the baseline (2010/2011). The reason for this apparent contradiction is the sizeable shift in enrollment toward the charter sector, which we discuss below.

In terms of ELA achievement growth, in 2016 the Newark district schools were ahead of where they were in 2010/2011, relative to comparable schools in New Jersey. In contrast to having achievement growth of 0.014 SD below average, Newark district schools had achievement growth of 0.044 SD above average ($-0.014 + 0.058$) in English. Meanwhile, Newark charter schools largely maintained their advantage in English, generating 0.168 SD more growth than did schools working with similar students in New Jersey ($0.215 - 0.047$).

Understanding the 2015 Increase

The timing of the sharp rise in Newark's achievement growth in 2015 corresponded with two other events that could have affected Newark's measured performance: the transition from the NJASK to the PARCC assessment and an increase in the proportion of students with no reported score on the state tests. We investigate both possible factors.

The PARCC differed from the previous NJASK assessment in both subjects. In math, the new tests placed greater emphasis on mathematical reasoning over procedural knowledge. And in ELA, the new tests placed more emphasis on written responses that employed critical thinking skills to analyze reading passages (NJDOE 2016). Moreover, the PARCC test in New Jersey was almost exclusively administered by computer (in contrast to the paper-based NJASK format). It is possible that PARCC assessed different content than did NJASK, and that Newark students may have performed better on this content even in the years prior to 2015, had they been tested on it. Although we have no way to test this theory, it seems plausible given

that Newark schools began emphasizing the Common Core State Standards earlier than did other school districts in New Jersey.

At the student level, the relationship between a student's achievement score one year to the next did not change with the transition to the new test. A student's performance on the 2015 PARCC test was only slightly less correlated with her performance on the prior year test as it had been in 2014 (0.82 compared to 0.84 in math and 0.80 compared to 0.83 in English). In other words, a student's prior performance on NJASK was equally predictive of her PARCC achievement as it was of her subsequent performance on NJASK in prior years. Kane et al. (2016) reported similar findings regarding the year-to-year correlations in test scores in four states administering Common Core aligned assessments in 2015 (Massachusetts, New Mexico, Maryland, and Delaware).

A student's performance on any test reflects a combination of factors, however: those that are stable (such as their family background, test-taking behaviors, prior learning) as well as recent additions to their knowledge (their improved understanding of the new standards). Student-level correlations could be attributable to family background and prior learning—which we would expect to be stable—even if the school-based additions to their knowledge changed with the switch in assessment.

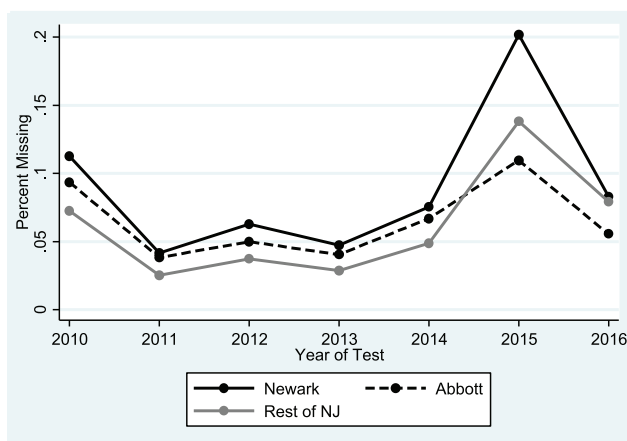
As a result, we estimated district-level differences in achievement growth within New Jersey in each of the years (which reflect additions to knowledge, after differencing out the effect of students' background and prior learning). During the NJASK years (2010–2014), the year-to-year correlation between district-level value-added averaged 0.64 for math and 0.66 for ELA.¹⁴ Even after excluding Newark, the correlation between district value-added in 2015 (with PARCC) and district value-added in 2014 (with NJASK) fell to 0.47 in math and 0.29 in ELA. Thus, the introduction of the PARCC test reshuffled district rankings in value-added more than in pre-PARCC years. Assuming both tests have similar levels of measurement error, these results imply that the PARCC and NJASK were assessing different sets of skills, and the districts that excelled in preparing students for PARCC were not necessarily the same as the districts that excelled at preparing students for NJASK. Thus, what appears to be a single-year gain in performance may have been present before 2015, but was simply undetected by earlier NJASK tests.

We also investigated whether the 2015 increase in Newark value-added was related to the increase in the proportion of students with no assessment results in the state data. Figure 2 reports the proportion of New Jersey students with missing math scores.¹⁵ Parent and teacher groups opposed to standardized testing (e.g., United Opt Out) were especially active in New York and New Jersey in spring 2015. Across New Jersey, the proportion of

¹⁴Correlations were similar among the largest quarter of New Jersey districts.

¹⁵The ELA plot was very similar.

Figure 2. Percentage of Students Missing Valid Math Test Scores over Time



Notes: Percentage of students in grades 4–8 missing math test scores in the academic years 2009–10 through 2015–16 in Newark, Abbott districts, and the rest of the state. The plot for English language arts (ELA) was similar.

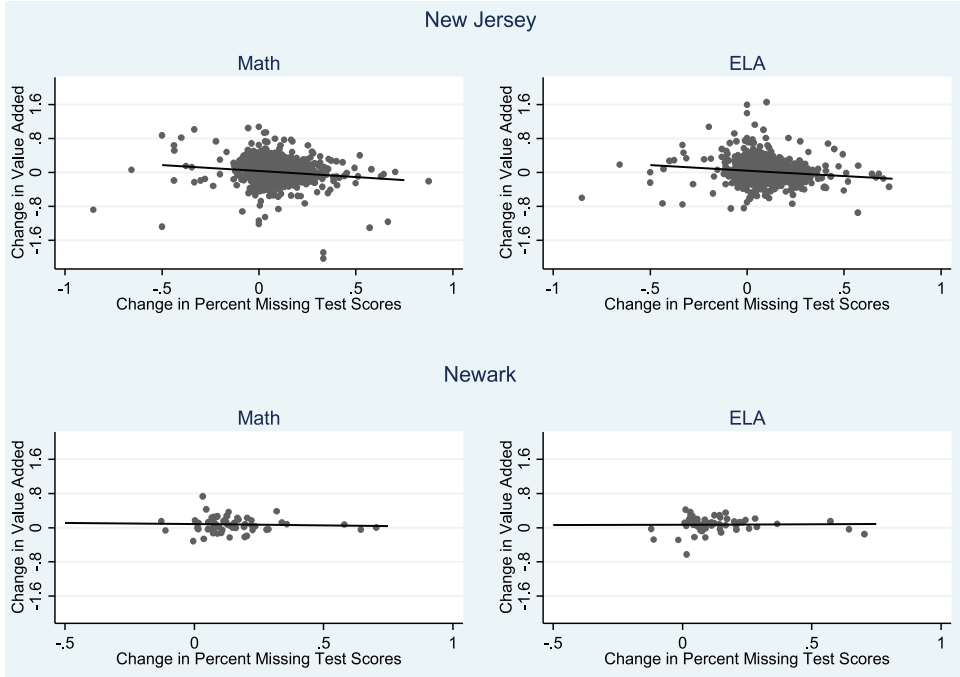
students without test scores spiked in 2015. Although we cannot distinguish between students who actively opted out and those lacking scores for other reasons, it is likely that the sharp increase was because students and parents opted-out of the PARCC test at greater rates (Ujjifusa 2015). The 2015 rise in the proportion of students missing scores was particularly large in Newark.

The increase in the proportion of students missing test scores does not seem to have caused the increase in Newark's 2015 student achievement, however. Figure 3 plots the change in school-level value-added against the increase in the proportion of students with missing scores between 2014 and 2015, separately for Newark and the remaining New Jersey districts. If the opt-out movement explained Newark's 2015 value-added increase, the schools with greater increases in the proportion of missing scores would experience greater value-added increases. However, the opposite seems to be true. The increase in the proportion of students missing scores was negatively related to the increase in a school's value-added, indicating that the rise in the proportion of missing test scores did not drive the increase in value-added in Newark.

Within-School Changes

Table 4 shows the change in value-added within existing Newark schools, both overall and separately, for three categories of schools: charter schools, the eight district schools renewed in 2013, and the remaining district schools. Columns (1) and (5) represent the overall estimates with the inclusion of school fixed effects to isolate within-school changes. In both math

Figure 3. Change in Percentage Missing Valid Test Scores and Change in Value-Added by School between 2014 and 2015



Notes: Changes in value-added and the percentage of students missing test scores are calculated by subtracting the 2013–14 academic year from the 2014–15 academic year. ELA, English language arts.

and ELA, continuing declines in achievement growth within schools through 2014 were evident. Charter schools had the largest within-school declines in achievement growth in both subjects in 2012 (columns (2) and (6)). There were especially large declines in renew schools in 2013 (the first year of renewal), when achievement growth declined by 0.17 SD in math and 0.13 SD in ELA relative to the baseline years. These declines were followed by a rebound in 2015 in both subjects. For ELA this rebound persisted into 2016, but within-school math value-added declined below pre-reform levels again in 2016. On a within-school basis, the average charter school achievement gain in math was still 0.19 SD below the pre-reform years in 2016. In the district schools that were neither renewed nor closed, achievement growth was 0.081 SD below pre-reform levels. By 2016, however, annual achievement growth in the renew schools had recovered to the point that they were not statistically different from pre-reform levels in either subject.

One possible explanation for the within-school declines in achievement growth is the disruption that occurred as new curricula were introduced, schools closed, their students were absorbed into other schools, and renew schools hired new teachers. In 2013, a spike in new students entering

Table 4. Within-School Changes in Value-Added in Newark (Including School Fixed Effects)

Change in Newark vs NJ difference relative to 2010/2011	Math				ELA			
	All (1)	Charter (2)	Renew (3)	Other (4)	All (5)	Charter (6)	Renew (7)	Other (8)
2012	-0.064*** (0.019)	-0.132*** (0.047)	-0.016 (0.023)	-0.057** (0.024)	-0.034* (0.017)	-0.082* (0.047)	-0.034*** (0.011)	-0.024 (0.021)
2013	-0.107*** (0.021)	-0.152*** (0.032)	-0.171*** (0.037)	-0.090*** (0.027)	-0.052*** (0.014)	-0.072*** (0.024)	-0.132*** (0.031)	-0.039** (0.017)
2014	-0.079*** (0.022)	-0.132*** (0.062)	-0.054 (0.054)	-0.076*** (0.024)	-0.091*** (0.015)	-0.126*** (0.034)	-0.108** (0.051)	-0.085*** (0.018)
2015	-0.002 (0.030)	-0.101 (0.082)	0.075*** (0.029)	0.015 (0.032)	0.007 (0.020)	-0.056 (0.055)	0.024 (0.034)	0.023 (0.020)
2016	-0.098*** (0.021)	-0.190*** (0.037)	-0.025 (0.033)	-0.081*** (0.027)	0.025 (0.019)	-0.049 (0.045)	0.035 (0.030)	0.049** (0.024)

Notes: Robust standard errors in parentheses (clustered at the school level). Estimates generated by a two-step process in which we first estimated a value-added model controlling for student demographics (indicators for race/ethnicity, gender, and free and reduced-price lunch status), prior achievement (including cubic polynomials of math and ELA achievement scores interacted with grade), peer covariates, district-by-year fixed effects, grade-by-year fixed effects, and interactions between our grade-by-year fixed effects and student and peer covariates. We then estimate a second equation in which the outcome is the residual from the first equation plus district-by-year fixed effects and the predictors are a main effect for Newark, Newark-by-year indicators, year fixed effects, and school fixed effects. Columns (2)–(4) and (6)–(8) each respectively are estimated based on a single regression in which we split up the main effect of enrollment in Newark into three categories: Newark charter schools, renew schools, and other schools. ELA, English language arts.

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 5. Regressing Growth in School Enrollment on School Math Value-Added

Years	Grades 4–8			Grades 1–3		
	Newark (1)	Abbott (2)	Rest of NJ (3)	Newark (4)	Abbott (5)	Rest of NJ (6)
2011 to 2012	55.45 (34.46)	29.03** (13.80)	0.668 (27.51)	47.66 (29.85)	18.84 (15.77)	–0.107 (11.83)
2012 to 2013	130.0** (51.43)	32.52 (32.15)	11.86 (22.40)	124.7** (51.65)	23.33 (32.07)	11.40 (13.88)
2013 to 2014	123.7*** (46.09)	6.794 (21.74)	27.63 (26.74)	61.33 (55.49)	17.04 (27.94)	14.48 (14.42)
2014 to 2015	122.8* (68.15)	22.79 (53.24)	4.136 (11.10)	174.5 (115.5)	–22.66 (25.94)	–15.68* (9.362)
2015 to 2016	165.1** (81.72)	46.50** (21.80)	14.31 (10.66)	29.68 (57.22)	25.57** (11.19)	–6.906 (7.685)
2011 to 2016	394.6** (187.3)	93.25 (58.46)	–27.70 (36.42)	405.1* (236.4)	–11.22 (40.55)	3.112 (18.65)

Notes: Estimates are coefficients from a school-level regression of change in enrollment on Empirical Bayes estimates of baseline value-added. Abbott and Rest of NJ regressions include district fixed effects. Standard errors in the Newark regressions are robust. Standard errors in the Abbott and Rest of NJ regressions are clustered at the district level.

* $p < .05$; ** $p < .01$; *** $p < .001$.

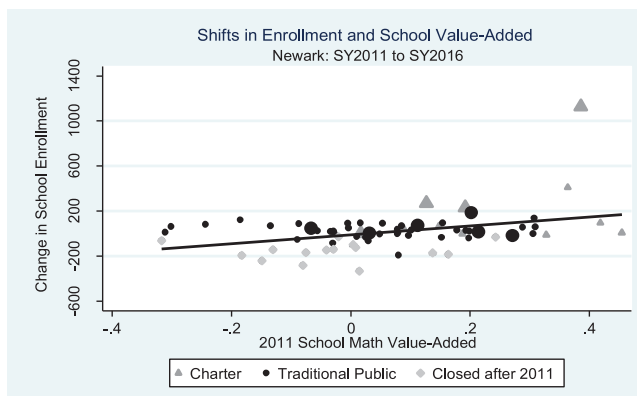
schools in Newark occurred, particularly in the renew schools that absorbed the lion's share of students from closed schools. This surge of new students may have contributed to the decline in value-added in both Newark's district and charter schools. In supplemental analyses not reported here, however, we control for the proportion of students in non-entry grades new to a school. This turnover did not explain the within-school decline in value-added.

In addition to student turnover, Newark's schools were experiencing significant teacher turnover during the early reform years. A surge in new teachers entered renew schools in 2013, many of whom were in their first year of teaching in the state. Specifically, nearly half of the renew school teachers were new to their school in 2013. Roughly a quarter of teachers were new in other district schools that year. This factor may help explain the initial decline in value-added in these schools. Moreover, we might have expected achievement to subsequently rise simply as new teachers gained experience.

Between-School Reallocation

In Table 5, we investigate the reallocation of students into higher value-added schools. In Newark, a significant positive relationship exists between a school's value-added and its subsequent enrollment growth. The relationship between value-added and enrollment growth holds in both the tested grades (4–8) and the younger grades (1–3). Looking over the entire

Figure 4. School Math 2011 Value-Added and Enrollment Changes from 2011 to 2016 in Newark



Notes: Enrollment changes calculated by subtracting 2010–11 enrollment from 2015–16 enrollment. Enrollment based on tested grades (i.e., 4–8). Larger markers indicate schools that were oversubscribed in 2015, according to the 2015–16 *Newark Enrolls Guidebook for Families*. SY, school year.

2011–2016 period, the estimates suggest that in Newark a school with 0.10 higher value-added in 2011 gained roughly 40 additional students.

Enrollment changes were much less related to differences in value-added in the Abbott districts and in the rest of New Jersey school districts. For these districts, no statistically significant relationship existed between enrollment growth and value-added differences in most years, and the magnitude of the coefficient in each year and over the whole period (2011 to 2016) was much smaller than in Newark.

Figure 4 plots enrollment growth in grades 4–8 from 2011 to 2016 against school math value-added in 2011 for all Newark schools. Substantial variation in value-added across schools in Newark is evident, ranging from below -0.30 to above 0.40 , and a clear upward slope indicates that higher value-added schools grew faster. The graph emphasizes that much of this relationship is attributable to the growth of charter schools and the closing of less effective district schools. Also, we highlight schools that were oversubscribed in the first year of universal choice (2015). Since most of these schools had higher value-added, there would have likely been even more reallocation toward high value-added schools had it not been for capacity constraints that prevented even larger enrollment shifts.

Decomposition

Table 6 decomposes the change in value-added from the pre-reform baseline years (2010–2011) to 2016 into within- and between-school components separately for Newark and the other Abbott districts. Between the baseline period (2010–2011) and the most recent available year (2016), average value-added in Newark grew by 0.07 SD in ELA—a sizeable gain comparable

Table 6. Decomposing the Change in Value-Added from 2010/2011 to 2016

	<i>Newark</i>		<i>Abbott</i>	
	<i>Math</i> (1)	<i>ELA</i> (2)	<i>Math</i> (3)	<i>ELA</i> (4)
Total change in value-added	-0.036 (0.007)	0.070 (0.007)	-0.014 (0.003)	0.026 (0.003)
Within school	-0.080 (0.006)	0.026 (0.006)	-0.021 (0.003)	0.018 (0.003)
Total between school	0.043 (0.003)	0.043 (0.003)	0.007 (0.001)	0.008 (0.001)
<i>Between existing schools</i>	0.049 (0.005)	0.038 (0.005)	0.002 (0.001)	-0.001 (0.001)
<i>Cross</i>	-0.017 (0.005)	-0.004 (0.006)	0.001 (0.001)	0.002 (0.001)
<i>School entry</i>	-0.001 (0.001)	0.002 (0.001)	-0.001 (0.000)	0.004 (0.000)
<i>School closing</i>	0.013 (0.002)	0.008 (0.002)	0.005 (0.001)	0.003 (0.001)

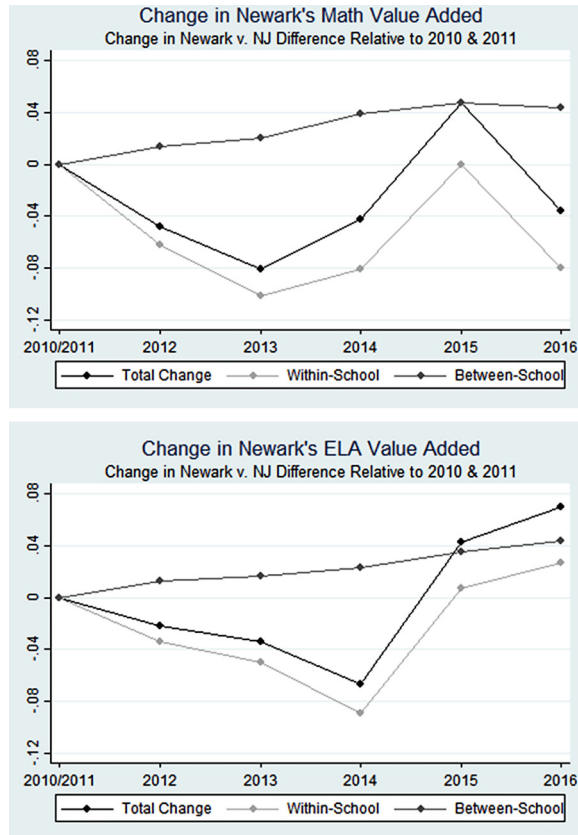
Notes: Change in average value-added of the schools attended by Newark students broken down into two components: improvements (or declines) in value-added within existing schools and changes due to the movement of students between schools with differing value-added. Change in value-added due to movement between schools is further broken down into four components: enrollment shifts between schools that existed throughout 2011–2016, the opening of new schools (school entry), the closing of schools, and a final component caused by a cross-product of changes in enrollment by changes in value-added. See text for further details. ELA, English language arts.

to the impact of being assigned to an experienced rather than novice teacher (Gordon, Kane, and Staiger 2006). However, 62% of the difference in ELA was attributable to the reallocation of students from lower to higher value-added schools.

In mathematics, the average achievement gain was largely unchanged in Newark, at -0.036 SD. Improvements resulting from the reallocation of students from low to high value-added schools (0.043) essentially offset a within-school decline of a larger magnitude (-0.08). In other words, Newark's math value-added would have declined if not for the shift in enrollment toward higher achievement growth schools. The Abbott districts saw negligible gains from between school movements: 0.007 in math and 0.008 in English (both statistically significant but small).

To illustrate the importance of the reallocation of students in Newark, we plot in Figure 5 the annual estimates of the cumulative change in value-added attributable to the within-school and between-school components. Although improvements due to between-school reallocations have grown steadily (to about 0.043 SD per year in both subjects), a U-shaped pattern of within-school changes emerges, with substantial initial declines followed by a large increase in 2015. In math, the increase was followed by a large decline in 2016, but in ELA, within-school improvements continued into the most recent year.

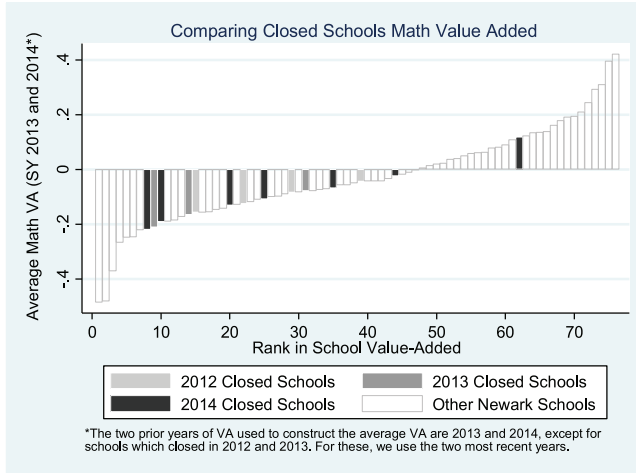
Figure 5. Decomposing Newark's Change in Value-Added Relative to New Jersey in 2010/2011



Notes: Plots the decomposition of Newark's change in value-added, relative to New Jersey in the baseline period, into its within- and between-school components using a version of Equation (4) in which instead of estimating a "long difference" between 2009–10/2010–11 and 2015–16, we estimate the year-to-year change. ELA, English language arts.

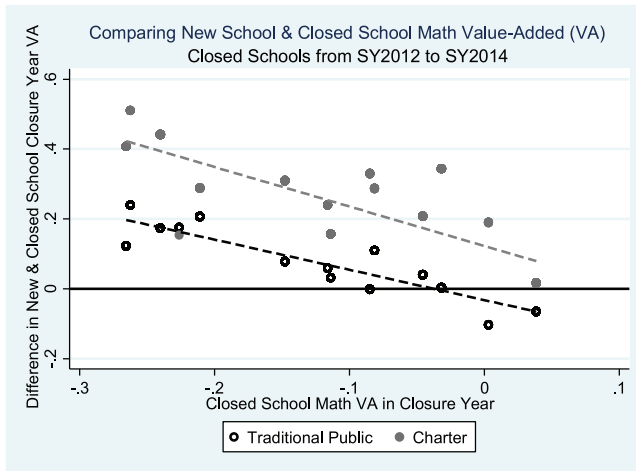
Most of the between-school improvement in value-added was driven by enrollment shifts between existing schools. Our decomposition suggests that school closures were an important part of the between-school improvement in Newark, simply because many closed schools had below-average value-added prior to closure. Figure 6 ranks all Newark schools by their math value-added, and highlights the K–8 schools that closed between 2012 and 2014. Closed schools tended to rank lower in value-added, although some schools that were not closed ranked even lower. Figure 7 plots the difference in school value-added between a closed school (in the year of closure) and the school the student moved to for all students in closed schools between 2012 and 2014. Students in closed schools with the lowest value-added moved to schools with substantially higher value-added, and those in closed schools with relatively high value-added saw little change in their

Figure 6. Math Value-Added Rank of Closed Schools Relative to Other Newark Schools



Notes: Average math value-added for all schools in the 2012–13 and 2013–14 academic years in rank order. The exception (*) is schools that closed in 2011–12 and 2012–13 for which we take the average value-added from the last two years they were open. VA, value-added; SY, school year.

Figure 7. Change in School Math Value-Added for Students Leaving Closed Newark Schools



Notes: For each student leaving a closed school, we calculate the difference in school-level math value-added between her closed school and new school using value-added calculated in the closure year for both schools (e.g., 2011–12 value-added for a school that closed at the end of spring 2012). We plot the average difference for students who moved to a district school (transparent) and those who moved to a charter school (dark gray) on the y-axis against the closed school’s closure year math value-added. Dashed lines (in dark gray and black) are fitted regressions for students who move to a charter school and district school, respectively. VA, value-added; SY, school year.

school's value-added. Additionally, students who moved to charter schools (approximately 20% of the closed school students) saw consistently larger gains in value-added, regardless of their closed school's value-added.

Theoretically, the students from the closed schools would not necessarily have experienced the same higher achievement growth as their new classmates. We explore this question in the Appendix and find evidence that closed school students do, on average, experience higher achievement growth after changing schools.

Conclusion

Although much has been written about the political controversies surrounding the Newark reforms, this article is the first systematic accounting of their impact on student achievement growth. By 2016, relative to the rest of the state, average student growth in Newark improved significantly in English and was not significantly changed in math. Our findings also highlight the difficulty of turning around existing schools, as reflected in substantial declines in achievement growth within both charter and district schools in the early years of reform. Although some examples of district turnaround yielding immediate academic improvements do exist (such as New Orleans, Louisiana [Harris and Larsen 2016] and Lawrence, Massachusetts [Schueler et al. 2016]), it is not uncommon for districts undergoing reform to see initial declines in student achievement growth (e.g., the Tennessee Achievement School District [Zimmer et al. 2017]). For Newark, the within-school rebound in achievement growth between 2014 and 2016 suggests that the tide may have turned (particularly in English).

As noted in Figure 1, Newark was not the only low-income district in New Jersey that witnessed an increase in achievement growth in 2015 and 2016. To the extent that the gains reflect policy shifts that Newark and other districts adopted, the improvement is encouraging. Because the rise coincided with the introduction of a new assessment, however, the change could be an artifact of measurement—a change in test scaling or testing procedures that disproportionately benefited students attending high poverty schools. We have tested a number of possibilities: the effect of accommodation policies on students with disabilities; floor or ceiling effects on the PARCC; a boost in scores for English language learners created by better text-to-speech options on the computer-based exams; and changes in the treatment of incomplete items in the scoring of NJASK and PARCC. None of these factors can account for the rise in achievement by many low-income districts in New Jersey. Nonetheless, even if a skeptic were to attribute the 0.07 gain in ELA achievement in the Abbott districts between 2014 and 2016 to some unspecified measurement artifact that also benefited Newark, the change in achievement growth in Newark was still twice as large as the other Abbott districts.

Still, the most distinctive element of the Newark reforms started in their first year (before the change in assessments) and continued through 2016: the steady shift in student enrollment toward higher value-added schools (both charter and district schools). That shift was helped along by the closure of some of the district's lowest value-added schools—which essentially forced parents to choose from a set of schools with, on average, higher value-added scores (there may be room for further improvement using this strategy given that some of the lowest value-added schools remained open). Steady expansion of the charter sector also occurred over this time period, with an emphasis on expanding enrollment at the high value-added charters. As a result, even in ELA, where we observe a sharp increase in achievement growth within existing schools, more than 60% of the improvement was attributable to between-school reallocation of students. And, in math, only after factoring in the between-school movement, which offset a within-school loss in achievement growth, did overall achievement growth remain on par with the district's relative performance in 2010/2011.

The experience in Newark has shown that reallocation of market share can be an important contributor to productivity growth in K–12 education, as it has been in many other industries. Nevertheless, other cities considering Newark-style reforms should remember two facts that make Newark unusual. First, reallocation was not always “market-based,” as school closures required some students to move involuntarily. Closing schools is politically difficult and may come with other downsides. Other districts, not under state control, may be less successful in making such hard choices. Second, Newark started the reform process with access to unusually effective charter schools nearby, which are not present in most other cities. It would be an oversimplification to suggest that most of Newark's progress was due solely to greater parental choice. In Newark, the positive effects of parental choice were enhanced by a series of difficult, but generally well-targeted school closures and ready access to an unusually effective charter sector.

Appendix

School Closure Effects on Student Achievement

Students in Newark schools that closed (i.e., “closed-school students”) subsequently attended schools with higher average value-added. In other words, students moved to schools where the *average* student was making faster progress. In this section, we explicitly test whether students from closed schools themselves subsequently outpaced similar students elsewhere in New Jersey.

If student learning was not disrupted when a student moved between schools, value-added is a valid estimate of a school's average treatment effect and that effect applies to all subgroups of students. Under these conditions, we would expect the achievement of students from closed schools to rise by the difference between their new and old schools' value-added. Those assumptions may not be true, however. For instance, moving schools as a

result of a closure may generate short-term achievement disruptions. Second, schools absorbing closed-school students may not maintain the same productivity (e.g., they hire novice teachers to accommodate expansion, the influx of new closed-school students disrupts other students' learning, and so on). Third, if treatment effects are heterogeneous, students from closed schools may not experience the same gains as the average student at their new school.

We estimate the effect of closure on achievement growth for students in three cohorts of schools that closed at the end of the 2011–12, 2012–13, and 2013–14 school years (referred to by their spring years from here on). The sample is limited to students who have a valid test score and lagged test score in the year of closure. Together, this sample includes 14 schools and 2,216 students in grades 4–8 in the closure year (four schools in the 2012 cohort, three in 2013, and seven in 2014).

We use a difference-in-difference model to estimate the closure effect. The first difference is growth before versus after closure for students in closure cohorts. The second difference is the growth of students in the rest of the state over this same period. We use the following model:

$$Y_{it} = \alpha + \sum_{k=-3}^3 \beta_k s_{ik} + \delta_{gt} + \lambda_c + \varepsilon_{it}$$

Y_{it} are the same student-level value-added estimates calculated using the two-step procedure involving Equations (1) and (2) as the main outcome of interest (see footnote 11). The term δ_{gt} represents a set of grade-by-year fixed effects, and λ_c are dummy variables for the three closed school cohorts (the excluded group is the rest of the state). Term s_{ik} represents a set of dummy variables for whether a student is in a closed cohort and the student is k years from closure (e.g., -1 corresponds to one year before closure) while β_k are the coefficients of interest for each year relative to the closure year.

Table A.1 reports that students in closed schools had lower average value-added than did students across the rest of the state. In the main specification pooling cohorts in columns (1) and (5), students in closed schools had value-added between 0.10 and 0.15 SD below the state average in math and 0.10 to 0.17 SD below the state average in ELA, depending on the cohort.

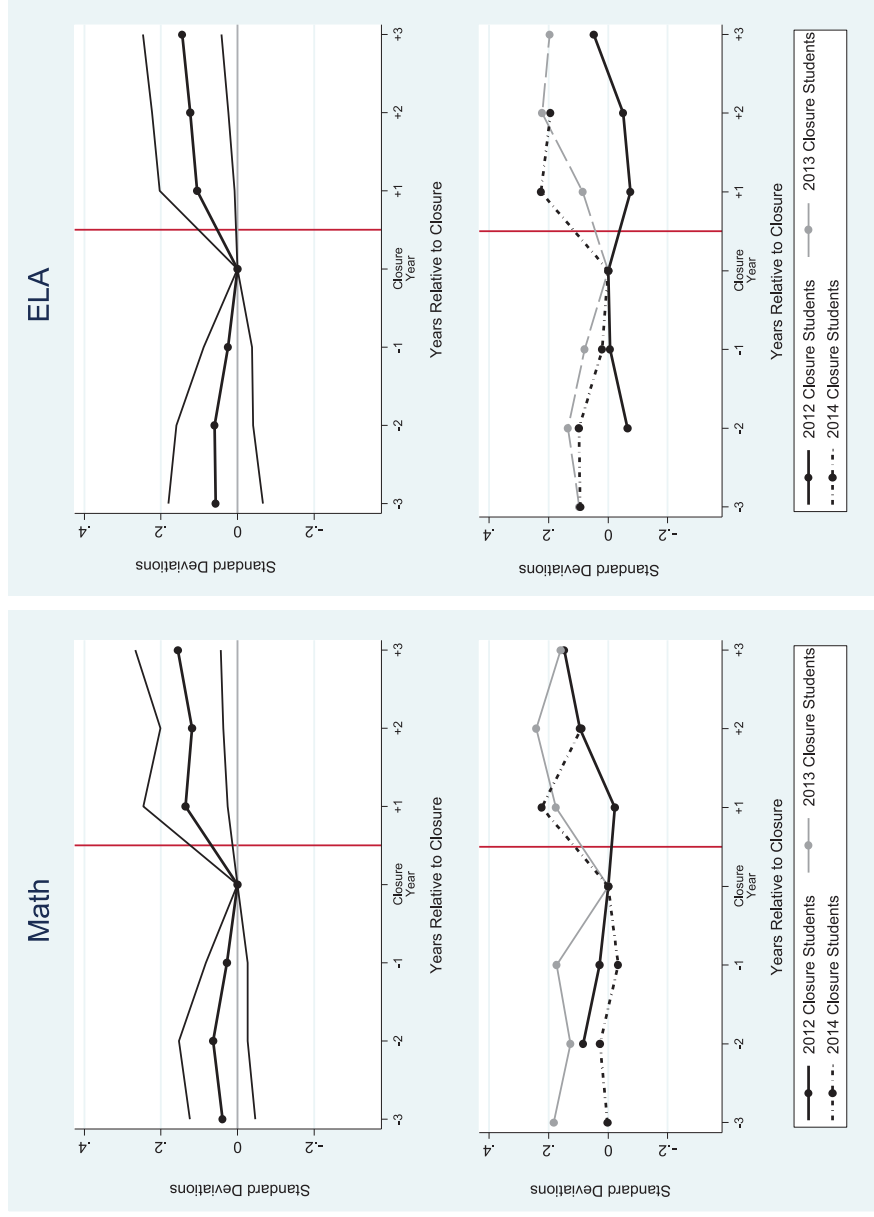
Also, there appear to be parallel trends in value-added on average between the closed school students and the rest of the state in the pre-closure period. None of the pre-closure coefficients using the pooled models are statistically different from zero in either subject. There is evidence of non-parallel trends for the 2013 cohort in the pre-period, however. In Figure A.1, this cohort of closures appears to have experienced a dip in value-added in the closure year, particularly in math. By contrast, the 2012 and 2014 cohorts had more stable pre-closure value-added trends on par with the state.

Table A.1. Changes in Value-Added for Students in Closed Schools in Newark

Change in closed schools vs NJ (relative to year of closure)	Math				ELA			
	Pooled (1)	2012 Cohort (2)	2013 Cohort (3)	2014 Cohort (4)	Pooled (5)	2012 Cohort (6)	2013 Cohort (7)	2014 Cohort (8)
3 Years before closure	0.039 (0.043)		0.182** (0.090)	0.002 (0.053)	0.057 (0.063)		0.097 (0.073)	0.094 (0.085)
2 Years before closure	0.063 (0.046)	0.083 (0.073)	0.126** (0.052)	0.027 (0.068)	0.060 (0.051)	-0.065 (0.096)	0.135*** (0.045)	0.098* (0.059)
1 Year before closure	0.028 (0.028)	0.029 (0.060)	0.172*** (0.047)	-0.034 (0.028)	0.025 (0.032)	-0.006 (0.068)	0.079* (0.042)	0.020 (0.041)
1 Year after closure	0.135**	-0.023 (0.074)	0.175*** (0.054)	0.223*** (0.077)	0.105** (0.050)	-0.075 (0.071)	0.085*** (0.036)	0.226*** (0.069)
2 Years after closure	0.119*** (0.042)	0.095 (0.058)	0.241*** (0.070)	0.090 (0.058)	0.124** (0.051)	-0.051 (0.069)	0.221*** (0.060)	0.193*** (0.070)
3 Years after closure	0.155*** (0.057)	0.147** (0.063)	0.160 (0.118)		0.145*** (0.052)	0.049 (0.071)	0.197*** (0.063)	
2012 Cohort	-0.150*** (0.034)	-0.118*** (0.039)			-0.167*** (0.043)	-0.077 (0.059)		
2013 Cohort	-0.102*** (0.036)		-0.172*** (0.033)		-0.098*** (0.034)		-0.133*** (0.020)	
2014 Cohort	-0.128*** (0.033)			-0.120*** (0.041)	-0.113*** (0.041)			-0.151*** (0.051)
Observations	3,238,680	3,233,664	3,233,069	3,235,075	3,230,593	3,225,562	3,224,976	3,226,989

Notes: Robust standard errors in parentheses clustered at the school level. Value-added for the outcome variable is calculated using the same two-step approach based on Equations (1) and (2) that was used in the decomposition (see footnote 11). All models include grade-by-year fixed effects. ELA, English language arts. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Figure A.1. Estimated Effects of School Closure on Student Value-Added



Notes: The top panels plot the estimated effects of school closure and 95% confidence intervals from the pooled sample results (columns (1) and (5) of Table A.1). The bottom panels plot estimated effects for each closure cohort (columns (2)–(4) and (6)–(8) from Table A.1).

Post-closure, we find that students from closed schools had a statistically significant increase in value-added in the year following closure in math and ELA. Specifically, these students grew 0.14 SD in math and 0.11 SD in ELA more in the first years after closure than they did pre-closure, controlling for statewide changes in growth over this period. This large gain in value-added relative to the closure year is consistent through the post-closure period in math and ELA, suggesting that closed-school students did reap the benefits of moving to higher value-added schools.

Figure A.1 plots the Table A.1 coefficients to show trends in value-added estimates for each year on average (top panels) and for each cohort (bottom panels).¹⁶ The 2013 and 2014 closure cohorts both experienced a jump in value-added one year after closure, increasing by 0.18 SD and 0.22 SD in math and 0.09 and 0.23 SD in ELA, respectively. In ELA, these gains were sustained in the post-closure period. In math, they dropped somewhat becoming statistically insignificant but remaining positive by the last year estimated for each cohort. The 2012 cohort fared worse than the other two cohorts in the years following closure. Students' value-added was no better off within two years after closure. By three years after closure, however, the 2012 cohort had gained 0.15 SD in math relative to change in the rest of the state.

Overall, students in Newark's closed schools appeared to benefit academically from closure, in both math and ELA. This growth was sustained for three years after closure on average, suggesting that schools that received closed-school students were able to transfer their higher average value-added to new students. The effects vary across cohorts: Students in the 2013 and 2014 cohorts experienced immediate value-added gains as a result of closure and those in the 2012 cohort did not initially. Perhaps of import, students in the 2012 cohort of closed schools were combined with students in existing schools that were targeted for renewal and were undergoing extensive reforms, which could explain the difference in the 2012 closure cohort's outcomes.

References

- Abdulkadiroglu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2011. Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics* 126(2): 699–748.
- Abdulkadiroglu, Atila, Joshua D. Angrist, Peter D. Hull, and Parag A. Pathak. 2016. Charters without lotteries: Testing takeovers in New Orleans and Boston. *American Economic Review* 106(7): 1878–920.
- Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak, and Christopher R. Walters. 2010. Inputs and impacts in charter schools: KIPP Lynn. *American Economic Review: Papers and Proceedings* 100(2): 239–43.
- . 2012. Who benefits from KIPP? *Journal of Policy Analysis and Management* 31(4): 837–60.

¹⁶Some cohorts have more post- or pre-closure years than do others because our data are limited to the 2010 to 2016 window.

- Baily, Martin N., Charles Hulten, and David Campbell. 1992. Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity, Microeconomics*, pp. 187–267.
- Baker, Bruce D., and Mark Weber. 2015. Research note: On average, are children in Newark doing better? *New Jersey Education Policy Forum*. Accessed at https://schoolfinance101.files.wordpress.com/2015/11/baker-weber-newarkbetteroff-njepf-11_15_15.pdf.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review* 103(1): 305–34.
- Brummet, Quentin. 2014. The effect of school closings on student achievement. *Journal of Public Economics* 119(C): 108–24.
- Calefati, Jessica. 2012. Principals selected for 8 struggling schools in Newark. *The Star Ledger*. Accessed at http://www.nj.com/news/index.ssf/2012/05/principals_selected_for_8_stru.html.
- Carlson, Deven, and Stephane Lavertu. 2016. Charter school closure and student achievement: Evidence from Ohio. *Journal of Urban Economics* 95: 31–48.
- Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. Healthcare exceptionalism? Performance and allocation in the U.S. health care sector. *American Economic Review* 106(8): 2110–44.
- 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): 2593–632.
- [CREDO] Center for Research on Education Outcomes. 2015. Urban charter school study report on 41 regions. Accessed at <https://urbancharters.stanford.edu/download/Urban%20Charter%20School%20Study%20Report%20on%2041%20Regions.pdf>.
- Conover, William J. 1999. *Practical Nonparametric Statistics*, volume 3. New York: Wiley.
- Dee, Thomas. 2012. School turnarounds: Evidence from the 2009 stimulus. NBER Working Paper No. 17990. Cambridge, MA: National Bureau of Economic Research.
- de la Torre, Marisa, and Julia Gwynne. 2009. When schools close: Effects on displaced students in Chicago Public Schools. Research Report. Chicago: Consortium on Chicago School Research.
- Engberg, John, Brian Gill, Gema Zamarro, and Ron Zimmer. 2012. Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics* 71(2): 189–203.
- Foster, Lucia, John C. Haltiwanger, and C. J. Krizan. 2001. Aggregate productivity growth: Lessons from microeconomic evidence. In Charles R. Hulten, Edwin R. Dean, and Michael J. Harper (Eds.), *New Developments in Productivity Analysis*, pp. 303–72. Cambridge, MA: National Bureau of Economic Research.
- Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98(1): 394–425.
- Fryer, Roland. 2014. Injecting charter school best practices into traditional public schools: Evidence from field experiments. *Quarterly Journal of Economics* 129: 1355–1407.
- Fulbeck, Eleanor, Martyna Citkowicz, Candace Hester, David Manzeske, Melissa Yisak, and Ryan Eisner. 2016. Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report. Washington, DC: American Institutes for Research.
- Gill, Brian, Ron Zimmer, Jolley Christman, and Suzanne Blanc. 2007. State takeover, school restructuring, private management, and student achievement in Philadelphia. Santa Monica, CA: RAND Corporation and Research for Action.
- Gordon, Robert, Thomas J. Kane, and Douglas O. Staiger. 2006. Identifying effective teachers based on performance on the job. The Hamilton Project, Discussion Paper 2006-01. Washington, DC: Brookings Institution.
- Harris, Douglas N., and Matthew F. Larsen. 2016. The effects of the New Orleans post-Katrina school reforms on student academic outcomes. Technical Report. New Orleans: Education Research Alliance for New Orleans.
- Heissel, Jennifer A., and Helen F. Ladd. 2016. School turnaround in North Carolina: A regression discontinuity analysis. *Economics of Education Review* 62: 302–20.

- Hess, Frederick M. 2007. Adequacy judgments and school reform. In Martin R. West and Paul P. Peterson (Eds.), *School Money Trials*, pp. 159–94. Washington, DC: Brookings Institution Press.
- Kamenetz, Anya. 2013. Cami Anderson and Zuckerberg, the unlikely duo behind Newark schools' revitalization. *Fast Company*. Accessed at <http://www.fastcompany.com/3007765/cami-anderson-and-zuckerberg-unlikely-duo-behind-newark-schools-revitalization>.
- Kane, Thomas J., and Douglas O. Staiger. 2008. Estimating teacher impacts on student achievement: An experimental evaluation. NBER Working Paper No. 14607. Cambridge, MA: National Bureau of Economic Research.
- Kane, Thomas J., Antoniya M. Owens, William H. Marinell, Daniel R. C. Thal, and Douglas O. Staiger. 2016. Teaching higher: Educators' perspectives on Common Core implementation. Center for Education Policy Research Report. Cambridge, MA: Harvard University.
- Kemple, James. 2015. High school closures in New York City: Impacts on students' academic outcomes, attendance, and mobility. New York: Research Alliance for New York City Schools.
- McGlone, Peggy. 2013. Newark school restructuring includes plans to put charters in district buildings. *NJ.com*. Accessed at http://www.nj.com/education/2013/12/newark_school_plan_restructure.html.
- Meyer, Peter. 2013. Newark's superintendent rolls up her sleeves and gets to work. *Education Next* 13(1): 2–5.
- Mooney, John. 2013. Fine print: Newark charter school revocation. *NJ Spotlight*. Accessed at <http://www.njspotlight.com/stories/13/04/15/fine-print-newark-charter-school-revocation/>.
- Morel, Domingo. 2018. *Takeover: Race, Education, and American Democracy*. New York: Oxford University Press.
- [NJDOE] New Jersey Department of Education. 2016. Historical context: Overview of New Jersey's statewide testing program. Accessed at <http://www.nj.gov/education/assessment/history.shtml>.
- [NPS] Newark Public Schools. 2013. Race to the Top District Application for Funding.
- . 2016. Newark Enrolls: frequently asked questions: School matches. Accessed at <http://newarkenrolls.org/frequently-asked-questions-school-matches/>.
- Papay, John. 2015. The effects of school turnaround strategies in Massachusetts. Paper presentation at the Association of Public Policy and Management. Miami, Florida.
- Russakoff, Dale. 2015. *The Prize: Who's in Charge of America's Schools?* New York: Houghton Mifflin Harcourt.
- Schueler, Beth E., Joshua S. Goodman, and David J. Deming. 2017. Can states take over and turn around school districts? Evidence from Lawrence, Massachusetts. *Educational Evaluation and Policy Analysis* 39(2): 311–32.
- Strunk, Katherine O., Julie A. Marsh, Ayesha K. Hashim, Susan Bush-Mecenas, and Tracey Weinstein. 2016. The impact of turnaround reform on student outcomes: Evidence and insights from the Los Angeles Unified School District. *Education Finance and Policy* 11(3): 251–82.
- Ujifusa, Andrew. 2015. "Opt-out" push gains traction amid Common-Core testing. *Education Week* 34(29): 16–17.
- Young, Viki M., Daniel C. Humphrey, Haiwen Wang, Kristin R. Bosetti, Lauren Cassidy, Marjorie E. Wechsler, Elizabeth Rivera, Samantha Murray, and Diane Whitmore Schanzenbach. 2009. Renaissance Schools Fund-supported schools: Early outcomes, challenges, and opportunities. Menlo Park, CA: SRI International.
- Zimmer, Ron, Gary T. Henry, and Adam Kho. 2017. The effects of school turnaround in Tennessee's achievement school district and innovation zones. *Educational Evaluation and Policy Analysis* 39(4): 670–96.

Copyright of ILR Review is the property of Sage Publications Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.