NEPC Review: School District Reform in Newark (National Bureau of Economic Research, October 2017) and Impact of the Newark Education Reforms (Center for Education Policy Research, Harvard University, September 2017)

Reviewers:

Mark Weber and Bruce D. Baker
Rutgers, the State University of New Jersey

December 2017

National Education Policy Center

School of Education, University of Colorado Boulder
Boulder, CO 80309-0249
(802) 383-0058
nepc.colorado.edu

Publication Type: NEPC Reviews are expert third-party reviews of selected non-peer-reviewed publications. Using academic peer review standards, reviewers consider the quality and defensibility of a report's assumptions, methods, findings, and recommendations. Written in non-academic language, reviews are intended to help policymakers, reporters, and others assess the social science merit of a reviewed report and to judge its value in guiding policy.

Funding: This review is one of a series made possible in part by funding from the Great Lakes Center for Education Research and Practice.

Suggested Citation: Weber, M. & Baker, B.D. (2017). *NEPC Review: "School District Reform in Newark (National Bureau of Economic Research, October 2017)* and *Impact of the Newark Education Reforms (Center for Education Policy Research, Harvard University, September 2017)."* Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/review-newark-reform.

NEPC Review: School District Reform in Newark (National Bureau of Economic Research, October 2017) and Impact of the Newark Education Reforms (Center for Education Policy Research, Harvard University, September 2017)

Reviewers:

Mark Weber and Bruce D. Baker
Rutgers, the State University of New Jersey

December 2017

Summary of Review

In 2010, Facebook CEO Mark Zuckerberg announced, with great fanfare, a $100 million challenge grant for the support of a series of reforms in Newark, NJ schools. The two reports reviewed here are the first attempt at a comprehensive assessment of the impact of the Zuckerberg donation on student achievement. As such, the reports, which were written by a team of economists, have garnered significant attention. In this review, we evaluate the reports and their potential contribution to education policy research.

- The reports find no increase in student growth or "value-added" in math and only nominal increases in English language arts over the five-year period following the grant. We explain that these small gains in English are most likely due to a change in assessments, rather than to any policies connected to the Zuckerberg donation. We also note that many districts close to Newark with similar demographics experienced similar gains in ELA relative to the rest of the state, calling into question whether the gains have anything to do with policy changes in Newark.

- The reports focus on a series of "reforms" purportedly initiated by the Zuckerberg grant. These reforms are divided into "within-school" (personnel changes, Common Core implementation, turnaround schools, and a teacher contract featuring differentiated pay) and "between-school" (school closures, charter school expansion, and universal enrollment) components. There is little evidence presented,

however, as to how these reforms were actually implemented or how they differed from other New Jersey schools, making any claim of a causal connection between the grant, the reforms, and student achievement growth suspect.

- The reports contend that the majority of the small gain in English was due to "between-school" reforms: students moving from less productive to more productive schools – specifically, to charter schools. We find, however, that the reports did not account for critical differences between Newark's district schools and charter schools. Key differences include resources, student characteristics, discipline, student attrition, staffing, and curricular narrowing. Assuming that the "between schools" locus is correct, the underlying change might be "charterness" or might be one or more of these related differences that have little to do with charterness. The results, therefore, are rendered inconclusive and provide no evidence in favor of the Newark-Zuckerberg reforms or the efficacy of moving students in urban districts to charter schools.

- The reports repeatedly claim to be a "productivity" analysis; however, they make no attempt to account for differences in school "inputs" – the resource differences that can have a profound effect on student achievement. Because Newark charter schools enjoy significant resource advantages over district schools, omitting those advantages from the analyses greatly diminishes the value of these reports for shaping the education policies of Newark schools.

- To summarize: the reports do not clearly define the treatment in question, omit important factors we know affect student learning and test score outcomes, are hampered by the use of crude data, and find what can, at best, be described as isolated and small effect sizes. Consequently, they provide little evidence in favor of the Zuckerberg-funded reforms, particularly when considering the documented disruption around Newark's schools that has occurred since 2010.

NEPC
NATIONAL EDUCATION POLICY CENTER

NEPC Review: School District Reform in Newark (National Bureau of Economic Research, October 2017) and Impact of the Newark Education Reforms (Center for Education Policy Research, Harvard University, September 2017)

Reviewers:

Mark Weber and Bruce D. Baker
Rutgers, the State University of New Jersey

December 2017

I. Introduction

In 2010, Facebook CEO and founder Mark Zuckerberg announced a challenge grant of $100 million to support Newark, New Jersey public schools. Over the next several years, nearly $200 million was disbursed with the intent of effecting a series of reforms in Newark with the stated goal of creating a national model of education reform.

In the fall 2017, two related reports were published by a team of economists which purport to evaluate the effects of these reforms. *Assessing the Impact Of the Newark Education Reforms: The Role of Within-School Improvement vs. Between-School Shifts in Enrollment* is published by The Center for Education Policy Research (CEPR) at Harvard University. *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* is published as a preliminary draft by the National Bureau of Economic Research.

Both reports were authored by the same research team: Mark Chin, Thomas J. Kane, Whitney Kozakowski, and Beth E. Schueler of Harvard University; and Douglas O. Staiger of Dartmouth College.

In both reports, the authors acknowledge their research is funded by the *Startup: Education Foundation,* which is now the Chan-Zuckerberg Initiative. The amount of funding the authors received is not disclosed.

These reports are the first attempt at a comprehensive assessment of the impact of the high-profile Zuckerberg donation on student achievement. As such, they have garnered significant attention.[1] In this review, we evaluate the reports and their potential contribution to education policy research.

II. Findings and Conclusions of the Report

The policy brief published by the Center for Education Policy Research (CEPR) is based on the same research as the working paper published by the National Bureau for Economic Research (NBER). The CEPR brief has less technical information than the NBER report, but both are clearly designed to make the case that a set of reforms, initiated in the 2011-12 school year in Newark, New Jersey, has led to substantial "achievement growth" for students in the district compared to similar students and schools throughout the state. Because the two publications are so closely linked, we refer to them collectively throughout this review as "the reports."

The reports claim the origin of the reforms in question is a well-publicized donation from Mark Zuckerberg, co-founder and CEO of Facebook, in support of education in Newark. Zuckerberg announced the donation September 24, 2010 on *The Oprah Winfrey Show*; he would provide $100 million in a challenge grant that would eventually be matched by other donors.[2] The $200 million, in total, represented approximately 4% of the Newark Public Schools' (NPS) budget over the five years of the grant.[3] In this review, we refer to the entire sum of money as the "Zuckerberg donation," including the funds raised as part of the challenge grant.

The reports detail the ostensible components of "reforms" that the authors claim were a direct result of the Zuckerberg donation: "... *it* [the donation and matching funds] *provided the city and the district with the flexibility to implement an ambitious slate of reforms.*"[4] The reports classify these policies as either "within-school" or "between-school" reforms. Table 1 summarizes how the reports place the reforms into these two categories.

Table 1

Treatment Component	Category
Personnel changes	Within
Common Core implementation	Within
"Renew Schools" and other turnaround efforts	Within
School closures	Between
Charter school expansion	Between
Universal enrollment	Between
Teacher contract	Within

The reports make the case that these reforms improved the productivity of the entire Newark publicly-financed school system. The improvements are measured using a value-added model (VAM), which attempts to hold constant factors such as student characteristics, prior test performance, student peer characteristics, grade-level effects, and year-to-year trends in test scores throughout New Jersey. In this way, the report claims to be able to evaluate the growth of Newark students' test outcomes relative to similar students throughout the state.

The report notes that Newark's schools had an advantage relative to the rest of New Jersey in math value-added prior to the reforms: 0.068 standard deviations (SD) in 2010 and 2011. There was no statistically significant advantage in English language arts (ELA) in the prior period. Charter schools in Newark showed a statistically significant advantage in both math (0.319 SD) and ELA (0.215 SD).

In all five of the years following the Zuckerberg donation, Newark's publicly-financed schools, both charter and NPS, showed either no change or a decline in growth in math using the reports' VAM methodology. Relative to 2010 and 2011, math value-added for the entire city was unchanged in 2016.

In ELA, Newark's schools showed a decline in growth relative to the 2010/2011 baseline from 2012 to 2014. In 2015, however, the aggregate growth for the city suddenly increased. By 2016, ELA value-added was statistically significant (0.070 SD).

To determine the cause of the changes (or lack of changes) in value-added, the reports use an econometric technique to "decompose" the effects found in the VAM models. Borrowing from research in other fields – notably, health care – the reports break down the change in value-added into "within" and "between" school effects: "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 toward more effective schools ("between-school" reforms)."[5] The reports conclude: "…62 percent of the difference in English was due to the reallocation of students from lower to higher value-added schools."[6]

The reports further find that the overall drop in math value-added (-0.036 SD[7]) is the result of a larger drop in within-school VA (-0.080 SD) offset by a smaller gain in between-school VA (0.043 SD): "In other words, Newark's math value-added would have declined if not for the shift in enrollment toward higher achievement growth schools."[8]

A key finding of the reports is that the largest increase in value-added, relative to other years, occurred in 2015. This year was notable for two reasons: first, the state switched from the former New Jersey Assessment of Skills and Knowledge (NJASK) to the Partnership for Assessment of Readiness for College and Careers (PARCC), a Common Core-aligned assessment in mathematics and ELA. Second, 2015 was the height of an "opt-out" movement, where many students refused to take the exams.

The reports claim that neither of these events can explain the sudden increase in value-added in that year. The reports do note that district-level correlations in value-added fell between 2014 and 2015, the year of the transition to PARCC.[9] The report asserts that this change "… implies 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." Despite this acknowledgement, the reports still assert that the positive change in ELA value-added is "educationally meaningful."[10]

The reports also dismiss the effect of the opt-outs by pointing out that, at the school level, increases in the percentage of students with missing exams between 2014 and 2015 correlate negatively to a school's increase in value-added for 2015 (the correlation is not formally reported but appears small, particularly in Newark). In other words: if opt-outs led to the

increase in value-added for 2015, we would expect to see a positive correlation between missing scores and value-added changes; the opposite appears to be true.

The reports conclude that parental choice in schools, "...enhanced by a series of difficult, but generally well-targeted school closures and ready access to an unusually effective charter sector,"[11] led to the gains in achievement growth. While urging caution in applying the findings to other contexts, the reports still assert: "The experience in Newark has shown that re-allocation of market share can be an important contributor to productivity growth in K-12 education, as it has been in many other industries."[12]

II. The Report's Rationale for Its Findings and Conclusions

The reports rely on econometric methods – some typical for the field of education policy, but some novel – to make causal claims regarding the Newark reforms. A two-stage VAM, similar to models used in assessing the effects of teachers on student achievement growth,[13] is employed to determine the size and statistical significance of the reforms.

Implicit in this method is the idea that the reforms in question are unique to Newark, unique to the time period studied, and have been clearly defined in this research. In other words: if the changes in growth are to be ascribed to the reforms, those reforms must be clearly described, isolated from other factors which may influence growth, and unique to Newark from 2012 to 2016.[14]

The reforms, which constitute the "treatment" in this framework, are decomposed into "within" and "between" school components. The method used finds its origins in economic research on industrial plants and hospital patient outcomes.

To determine the extent of any achievement growth found in the research, the reports rely on standardized test scores from two different statewide assessments: the NJASK (2010 to 2014) and the PARCC (2015 to 2016). Because the effects are aggregated across tested grade levels (grades 4 to 8[15]), and because two different assessments were used during the period studied, the reports "standardize" the scores: that is, convert them to a common scale. The reports claim this standardization is sufficient to aggregate the effects across both grade levels and year, regardless of the test that was administered.

The reports note that other low-income districts in New Jersey saw a substantial increase in growth in 2015, the first year of the PARCC. While the authors admit the increase may be due to an "artifact of measurement,"[16] they discount the possibility, claiming to have tested a number of alternative theories.[17] The reports, however, do not describe the methods employed or the results of their tests. Further, they claim the change in growth outstrips that of other low-income districts: "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 which also benefited Newark, the change in achievement growth in Newark was still twice as large as the other *Abbott* districts."[18]

The reports give effect sizes in standard deviations – typical for econometric research but difficult for the lay reader to understand. In an attempt to provide context, the reports compare the size of the effects in this study to the "…impact of being assigned to an experienced versus novice teacher."[19]

IV. The Report's Use of Research Literature

There is a large body of research literature, spanning many years, that relates to the components of the Newark "reforms" in question – what researchers in the social sciences would refer to as the "treatment" – as presented by the reports. While the reports cite research relevant for some of the components, the research base for others is largely ignored.

A new teacher contract in 2012, for example, is cited as a within-school reform. The contract included "…differentiated teacher compensation, including incentives for high performers to stay in low-performing schools."[20] Yet no specific studies on teacher compensation are cited and summarized as part of making the case that this "reform" could have an effect on student achievement growth.

Further, there is no attempt in the reports to summarize the research that is cited. The NBER report, for example, cites a list of eleven studies as part of a "growing literature on school and district turnarounds,"[21] yet it makes no attempt to build a theoretical framework from those studies by synthesizing their findings and conclusions.

In addition, the review of the literature for some components is, in some cases, quite narrow. Charter school expansion is repeatedly cited as an important component of the Newark reforms; however, the research on charters cited in the reports is almost entirely limited to studies of oversubscribed Massachusetts or New Orleans charters, or of the national charter management organization, KIPP.

The only exception is a reference to the 2015 CREDO urban charter school report,[22] referenced only to make the claim that "… Newark is home to one of the most effective charter sectors in the nation in terms of student growth on standardized exams."[23] No additional research is cited to support this claim; in fact, no other previous research on Newark's charter schools is cited. Further, with the exception of an evaluation of year one of the 2012 teacher's contract and a footnoted citation of a research note by the authors of this review, no research on Newark's schools of any type is referenced. We provide a list of such research in Appendix D.

The decomposition methodology is derived from several papers cited in the reports, most notably a study of hospital patient outcomes.[24] The other studies cited apply decomposition to manufacturing[25] and the auto repair sector.[26] The reports cite no previous use of the decomposition method in studies of educational productivity, and no explicit argument is made for the validity of using test score outcomes as the outcome variable in this method.

V. Review of the Report's Methods

This section of our review is divided into the following subsections:

- Defining the treatment: A discussion of the research framework of the reports and important context which must be included to properly evaluate the effect of the Zuckerberg donation on student achievement.
- Charter school expansion and universal enrollment: The two key components of the "between-school" reforms.
- Missing components of the treatment: Factors not discussed in the reports and how they may affect the outcomes measured.
- Effect size: How the reports describe the size of the effect they found, and their attempt to put that effect into context.
- The larger context: Viewing the effects of the Newark reforms in historical context.

Throughout this section, we refer to the Appendices of this review, which provide quantitative analysis and further context.

Defining the Treatment

We begin our critique of these reports' methods by noting the authors repeatedly refer to their work as "productivity" research (emphases ours):

- "We assess the degree to which this opportunity improved the system's overall *productivity*."[27]
- "By examining district-wide *productivity* change, our study incorporates both types of spillover effects."[28]
- "We 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."[29]
- "The experience in Newark has shown that re-allocation of market share can be an important contributor to *productivity* growth in K-12 education, as it has been in many other industries."[30]

While there may be subtle disagreements, the definition of productivity is generally agreed upon both within and beyond the field of economics: output per unit of *input*.[31] As Burkhead and Hannigan (1978) note: "At some level of abstraction, the economist's definition of productivity is extremely simple and straightforward. It depends on an input-output relationship in which factors of production – land, labor, and capital – are converted into outputs."[32]

There is a long history of research on productivity in American public schools. Much of this research attempts to estimate an "education production function," linking school inputs – for example, spending per student – to student outcomes – for example, test score changes.[33] The goal of such research is to determine the practices and organizational structures that allow schools to get more "bang for the buck."

The Newark reports are certainly concerned with outputs as measured by test scores; however, missing from their methods are any accounting for variations in *inputs*. No measures of fiscal or other inputs are used within the value-added models, the decomposition methods, or the descriptive tables presented. For this reason, the reports cannot be considered a productivity analysis: they do not evaluate the measured outputs relative to any inputs.

In Appendix B, we explore what a true productivity analysis of Newark's schools might entail; we then conduct our own productivity analysis and show that, depending on the models and variables used, conventional wisdom about the relative productivity of Newark's schools can plausibly be challenged.

The omission of inputs within the analysis framework greatly limits the usefulness of these reports in informing the making of education policy. Any attempt to compare school outcomes across a wide variety of contexts should, to the fullest extent possible, account for variations – including inputs – that could impact those outcomes. Leaving inputs out of the analysis could lead readers to draw conclusions about the efficacy of the Newark reforms that attribute the cause of any achievement gains to improvements in school efficiency, rather than advantages in school resources.

There is little question that inputs such as spending per student exert significant influence on measured outcomes. A substantial and growing body of research finds that school funding has a significant and lasting effect on student achievement.[34] Unfortunately, we have shown in previous research that several New Jersey school districts are among some of the most financially disadvantaged districts in the nation.[35] Many of these districts are so-called "Abbott districts," and are part of the counterfactual group of schools in the reports' analysis. To the extent that the entire system of publicly-funded Newark schools enjoys a fiscal advantage over these districts, gains may be misattributed to the Newark reforms.

This issue is further exacerbated, as we describe below, by the growth of the charter school sector in Newark. The reports contend that the sector is "unusually effective"[36]; however, they could have just as easily noted the sector is unusually well-resourced. Moving students to charter schools that enjoy an input advantage may lead to greater achievement growth. The public policy ramifications, however, are very different than policies that move students to more efficient schools – schools that get better results with the same level of inputs.

In order to demonstrate how accounting for variations in inputs can affect a true productivity analysis, we undertake such an analysis in Appendix B of this review. While we are admittedly constrained by the data available – a constraint, we note, shared by the reports' authors – we nonetheless provide evidence that measures of the relative productivity of schools can be affected by including inputs in an analysis.

The reports' lack of attention to inputs is part of a larger issue: the lack of a clearly defined and measurable treatment. The reports attempt to show a causal connection between the Zuckerberg donation, the Newark "reforms," and achievement growth in Newark students. It is difficult enough to connect the donation to the reforms: it is possible that at least some of them, such as charter school expansion, could have been implemented in the absence of Zuckerberg's grant.

But as difficult as it is to tie the grant to the reforms, it is even more difficult to establish a causal link between the reforms and any achievement growth if a clear definition of the reforms is not established. This requires first delineating which of the reforms constitute the treatment and verifying that they were actually put into place. The components of the treatment then need to be measured to the extent that they can be. Finally, the researchers should establish whether those same reforms are present or not in the counterfactual school districts, and, again, measured as best as possible.

Table 2 lists the components of the Newark reforms that are described in the reports. We summarize here how the reports describe each of the components, and we note whether and how the authors attempted to validate their description of each component. We then draw upon our experience as long-time researchers of New Jersey education policy to add additional context to the reports' descriptions. We refer readers to Appendix A in particular, which provides crucial historical context for understanding education policy changes in Newark.

In addition, we describe several components of the treatment which were not included in the reports, but likely have a significant effect on the measured outcomes.

Our goal here is to provide a more complete understanding of the Newark reforms, allowing readers to make a better judgment as to whether the Zuckerberg funds did, in fact, have a direct effect on achievement gains by improving school productivity. Below the table, we describe each component and provide references to this review's Appendices, which contain further quantitative and other analysis.

Table 2

Reports' Description of the "Treatment"			
Treatment Component	**As Described by Authors**	**Author Sources**	**Missing Context**
New Teacher Contract	- New teacher evaluation system. - Differentiated teacher compensation. - Extended learning time in 28 schools. - Greater school-based decision making. - "...Newark was able to retain teachers with higher ratings at higher rates under the new contract."[37]	*Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research.[38]	- AIR report does not compare teacher retention rates before and after the contract. - "An arbitrator has found that **the state violated many of the terms of the 2012 contract agreement it signed with the Newark Teachers Union** – and which Gov. Chris Christie lauded at a news conference – and owes hundreds of Newark public school teachers millions of dollars."[39]
Differentiated Teacher Pay	- "Agreement reached on new pay-for-performance teacher contract."[40] - "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."[41]	*Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research.	- The actual amount of differentiated teacher pay has been reported to be far less than $20 million.

New School Leaders	"Anderson moved swiftly on her strategy to vest principals with much of the responsibility for reforming the district. She hired seventeen new ones in her first summer, recruiting from around the country, and within three years had replaced well over half of the seventy she had inherited."[42]	Russakoff, D.: *The Prize* (2016). (Note: the reports do not directly cite the quote given in this table.)	- The union representing NPS principals claims school leaders were punished for public comments against the Newark reforms, and not necessarily for their performance. - Russakoff notes the donation money set aside for the principal's contract was not spent.
Turnarounds ("Renew Schools")	- Rehired/replaced staff. - Extended learning time. - Professional development. - "Better access" to services.[43]	- Two newspaper reports and one magazine article. - AIR report. - Russakoff, D.: *The Prize* (2016).	No evidence is presented which validates the claims that teaching staff quality improved, learning environments improved, or access to support services increased.
Common Core	Newark emphasized CCSS alignment earlier than other districts.	No validation is offered.	- Common Core implementation is not, by itself, and indicator of instructional improvement. - There is no evidence NPS implemented the CCSS earlier than other NJ districts.

School Closures	Results find students in lower value-added schools moved to higher value-added schools.	Reports' findings through the decomposition methodology.	- Newark schools, especially charters, vary in resource allocation, which may account for differences in achievement growth. - NPS ratings of school characteristics were not reliable indicators of student achievement.
Charter Expansion	Reports characterize Newark's charter sector as "unusually effective."	Center for Research on Education Outcomes (2015). *Urban charter school study report on 41 regions*.	- Student population characteristics differ significantly between NPS and the charter sector. - Variables used in the CREDO studies (and the reports' VAM) are crude and likely do not capture significant student differences, leading to bias in the estimates of effects. - Newark charter schools' enjoy significant resource advantages over NPS.

Universal Choice	"Therefore, 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."[44]	No validation of the claim for a causal effect of universal choice on student achievement is presented.	- Most "popular" schools tended to be more racially segregated. - "One Newark" did not include all publicly-financed Newark schools. - 8% of students enrolled in Newark charter schools do not reside in Newark.

Missing and Relevant Components of the Report's Description of the "Treatment"	
Treatment Component	**Missing Context**
Funding Differences	Fiscal advantage to charters from: - Philanthropy. - Additional governmental support for facilities. - Staff differences leading to fiscal advantages. - "Hold harmless" state funding.
Curricular Narrowing	Shift in instruction toward tested subjects: - Shift to charters means less personnel in untested subjects. - Charter leader philosophies suggest an emphasis on tested subjects and test-taking strategies.
"Teaching to the Test"	Test gains due to assessment alignment are not necessarily indicative of instructional or curricular improvements.

The New Teacher Contract

As noted in the reports' timeline, a new teacher contract, including pay-for-performance incentives, was announced on October 18, 2012[45]. The reports cite an American Institutes for Research (AIR) brief[46] to make this claim about the impact of the contract: "…Newark was able to retain teachers with higher ratings at higher rates under the new contract."[47] In fact, the AIR brief only reports retention rates for the 2012-13 and 2013-14 years; there is no

evidence presented to validate the claim the retention rates rose after adoption of the new contract.

In the years following the original agreement, the Newark Teachers Union (NTU) has disputed whether Newark Public Schools (NPS) has adhered to the provisions of the contract. A series of grievances filed by the union culminated this fall in an arbitrator's decision[48] which found the districts had violated several provisions of the contract, including the terms of the teacher evaluation system cited in the reports.

It is impossible to determine whether the NTU contract had an effect on student achievement without first validating which provisions of that contract, if any, were actually implemented. Given the disputes over its implementation, it is incumbent on the authors to validate their descriptions of the contract's effects before making a causal claim.

Differentiated Teacher Pay

News stories from 2012 noted the provisions of the contract related to teacher compensation included up to $20 million for so-called "merit pay." One initial report claimed the amount of money available for merit pay could be as high as $80 million[49]; another early report put the figure at $18 million.[50]

The reports are ambiguous on the amount of merit pay that was actually distributed: "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."[51] According to the NTU Memorandum of Agreement[52], however, $20 million was available in "rewards" across the entire span of the contract, and not in the first year alone. Further, the reports never fully describe what the term "stipends" entails.

In any case, subsequent news stories noted the actual amount of incentive pay distributed was far less than original news stories had suggested it would be, and far less than described in the reports. *The Wall Street Journal*, for example, reported that the amount of "merit pay" distributed to 190 teachers in the first year totaled $1.4 million; the news story calculates only 5 percent of the total teaching force received any monies.[53] Other media reports state that $1.5 million was distributed to 233 teachers in the contract's second year.[54]

Without a clear, unambiguous accounting of how and how much differentiated teacher pay was distributed, it is not credible to make any claim this component of the Newark reforms had any effect on student achievement.

New School Leaders

One of the most controversial aspects of the Newark reforms has been the turnover of school leaders. In one case, union officials representing principals charged that NPS had suspended them for publicly voicing opposition to proposed reforms (NPS officials refuted the charge).[55]

The reports rely on two sources to make the case school leadership changes had an effect on student achievement: Dale Russakoff's *The Prize*, and an interview with NPS Superintendent Cami Anderson in *EducationNext* that can fairly be described as flattering.[56] Russakoff reports Anderson replaced many principals: "She hired seventeen new ones in her first summer, recruiting from around the country, and within three years had replaced well over half of the seventy she inherited."[57] But no evidence is presented that the quality of the principals improved.

In addition, there is no detailed description offered as to how the Zuckerberg donation helped to facilitate changes in principals; in fact, Russakoff notes that $13.7 million was committed to a principals' contract, but remained unspent.[58]

Turnarounds/"Renew" Schools

In 2012-13, NPS chose eight schools for "renewal"; the district chose another eight in 2013-14. The reports rely on five sources for a description of renewal; two are short newspaper reports that simply note "renewal" meant new principals and changes in teaching staffs.[59] The third is a magazine article that, once again, can be fairly described as flattering to Anderson. Its description of the Renew schools is limited: "...extra training for teachers, hiring bonuses for high-needs classes, more computers and Wi-Fi, and more access to social services such as nurses, social workers, and community mentoring."[60] There is no indication the author of the article confirmed any of this independently.

The fourth source is AIR's evaluation of the teachers' contract. In a footnote, it describes "renew" schools:

> "Renew schools are those schools identified for renewal based on a number of factors, including academic performance and enrollment over time, building utilization, and the age and condition of the building. Renew schools have a longer school day, increased professional development time for leaders and teachers, community organizing, and increased social and emotional supports."[61]

Only one of these factors is further expanded on in the AIR report: the implementation of extended school days and years. The report contends that this extended time could improve student learning. It also notes, however, that the NTU contract called for extra compensation to teachers who worked in extended-time schools.[62] Again, in a true productivity analysis, this difference in inputs would be accounted for.

The reports' last source on renew schools is, again, Russakoff's *The Prize*, which makes several bold claims:

> "There is no question that Anderson had equipped the Renew schools to serve children better. She had given them stronger principals and teachers, assessments to measure what students were and weren't learning, curriculum aligned to the Common Core standards, online learning programs that indi-

> vidualized instruction." (p. 187)
>
> "One problem was that the quality of teachers at the Renew schools was mixed." (p. 188)
>
> "Indeed, almost every aspect of the [Renew] schools had improved: stronger teachers and principals, more rigorous curriculum, longer school days, and noticeably better learning environments." (p. 203)

These are audacious claims; and yet, beyond a few anecdotes, there is no evidence – qualitative or quantitative – to support Russakoff's assertions. No evidence compares the effectiveness of teachers and principals before or after renewal. No analysis is provided of either the curriculum or the assessments mentioned. No description is given of the online programs, or how they individualized instruction.

While Russakoff's book may be an intriguing political read, it is simply not a sufficient source to provide a description of a treatment being studied for a potential causal effect. It does, however, confirm a finding from the reports: "renewal" did not lead to any substantial gains in value-added. As the reports note: "In the district schools that were neither renewed nor closed, achievement growth was 0.081 SD below pre-reform levels. However, by 2016, 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."[63]

Common Core Implementation

New Jersey is one of nine states that moved its statewide assessment program to the PARCC in 2015. PARCC is an assessment designed to align with the Common Core State Standards (CCSS), which were adopted by New Jersey in 2010[64] (in 2016, the state adopted a revised version of the standards[65]).

The reports repeatedly assert that Newark adopted the CCSS earlier than other New Jersey districts.[66] Yet no evidence is presented to back up this claim: while the reports claim Newark adopted aligned instructional materials in 2013-14, they do not present any evidence that this was early compared to other districts around the state.

The reports state that Expeditionary Learning, a curriculum adopted by the district in 2013-14, is "…highly rated by EdReports for Common Core alignment."[67] The EdReports reviews, however, only cover grades 6 to 8.[68] There is no evidence presented to suggest the district's math curriculum was better aligned to the CCSS. In addition, there is no indication the previous Newark curriculum was *less* aligned to the CCSS. It is also unclear in the report whether Newark's charter schools adopted this curriculum.

The reports cite a policy brief[69], written by two of the authors (among others), which surveys teachers and principals about CCSS implementation in five states; New Jersey is not one of them.[70] While limited due to its self-reporting design, the brief does provide some evidence that variations in instruction and organization can affect student outcomes on the PARCC.

What the brief does not -- indeed, cannot -- do is provide any evidence that Newark led the rest of the state in CCSS alignment.

And, as we discuss below, there is a further issue: whether the curricular and other changes that may have occurred are actual improvements in instruction, or simply gains due to students learning better test-taking strategies.

School Closures

The results of the decomposition method employed by the reports lead the authors to suggest that school closures were an important part of the effects on achievement growth. The reports note: "Students in closed schools with the lowest value-added moved to schools with substantially higher value-added, while those in closed schools with relatively high value-added saw little change in their school's value-added. Additionally, students who moved to charter schools saw consistently larger gains in value-added, regardless of their closed school's value-added."[71]

Again: because the reports' methodology does not take into account resource differences, it is impossible to say whether the gains in value-added were due to students moving to more efficient schools, or moving to schools that had resource (and perhaps other) advantages. This is particularly relevant, as we show below, to the strategy of closing district schools and placing more students into charters.

The reports also make this interesting observation: "Closed schools tended to rank lower in value-added, although some schools that were not closed ranked even lower."[72] This begs a question: was NPS pursuing a strategy of deliberately shuttering those schools that were the least effective?

Our previous research on Newark schools suggests the process was, at best, haphazard.[73] When the district released its ratings of schools in the first year of universal enrollment, our analysis showed the classifications of the schools were not reliable indicators of their students' achievement when controlling for student characteristics.[74]

Charter School Expansion and Universal Enrollment

Charter Schools: Student and Staff Characteristics

We set our discussion of charter schools and the aligned universal enrollment system here in a separate section, as charter expansion has been a crucial part of the Newark reforms and deserves a more thorough discussion. In Appendix C, we provide an in-depth analysis of state data regarding Newark charter school student populations, staff characteristics, and student cohort attrition. To summarize our findings:

- Newark's charter schools enroll significantly lower percentages of students with

a special education classification. The classified students the charters do enroll tend to have lower-cost and less profound learning disabilities.

- Historically, the largest charter networks – North Star and TEAM/KIPP – enroll proportionally fewer students who qualify for free lunch, an indicator of deeper economic disadvantage than reduced-price lunch.
- Year after year, North Star and TEAM/KIPP enroll no to very few English Language Learners (ELL).
- There is a significant cohort attrition at North Star and TEAM/KIPP; in other words, the size of student cohorts shrink as they pass from grade to grade, indicating that students leave the charter schools and are not replaced. This is particularly evident for black boys.
- North Star and TEAM/KIPP enroll larger proportions of female students than NPS schools.
- North Star and TEAM/KIPP have historically high suspension rates compared to NPS schools.
- Controlling for experience and degree, North Star and TEAM/KIPP pay more competitive wages than NPS schools, suggesting a resource advantage that may translate into longer school days and school years.
- North Star and TEAM/KIPP may be able to offer relatively higher wages because they employ many more inexperienced – and therefore, less expensive – teachers.

The reports do not acknowledge these differences between NPS schools and the city's largest charter schools. Instead, they place great emphasis on the claim that Newark's charters are "unusually effective."[75] The source for the claim is the 2015 CREDO report on urban charter schools, which asserts that Newark's and Boston's charter schools are the two highest performing sectors within the 41 urban areas studied.[76]

As we have noted in the past[77], the CREDO reports employ a virtual matching methodology that is only as good as the variables used to create matches to counterfactual students. This places a significant limitation on the interpretation of the effect sizes found, as the variables used to match students are crude, binary measures of student characteristics such as socio-economic status, special education status, and limited English proficient (LEP) status.

This is highly problematic when studying a school district like Newark, which enrolls a high percentage of economically disadvantaged students, LEP students, and special education students. In Newark, the vast majority of the student population qualifies for free or reduced-price lunch, the metric used to measure socio-economic status (SES). That does not mean, however, that the students don't vary, *relative to each other*, in SES. Further, there is reason to believe the distribution of students by SES is not equal at all schools.

This issue, which clouds the findings of the CREDO studies, is just as germane to the reports

reviewed here. The VAM employed relies on the same binary measures of student SES, LEP status, and special education status as found in the CREDO reports. And yet, as we show in Appendix C, we have evidence that suggests charter school students in Newark differ from students in the district schools in many significant ways – ways that have a direct impact on student achievement.

Universal Enrollment

In 2013, Newark implemented a universal choice system, called "One Newark," which allowed families to choose from both district and charter schools. As we note above, our previous research suggests NPS's rating of schools was not an accurate reflection of their student achievement, controlling for student characteristics.

Research suggests the interplay of parents' socio-economic status and their ability to navigate a choice system of schools is highly complex.[78] Whether universal enrollment impacts this interplay is an open question. Our research on the first year of "One Newark," however, suggests that the most popular schools in Newark were not necessarily those that achieved the highest outcomes on test scores after controlling for student characteristics. In addition, the most popular schools were among the most racially segregated in Newark.[79]

Several charter schools in Newark did not participate in the "One Newark" system.[80] If we grant the unproved assumption that universal enrollment makes charter schools more easily accessible to families with less social and other capital, it would follow that these schools were not equally accessible to all families within Newark, possibly creating and/or reflecting peer effects.

Finally, we note that, according to NJ state records, approximately 8 percent of Newark charter schools students in 2017 resided in districts other than Newark (see Appendix C). There is no indication in the reports that these students were excluded from the study group. This creates a potential problem with the decomposition methodology, as increased enrollments in charter schools are not due entirely to an increased share of resident Newark students. There is also the possibility that peer effects from charter school expansion are due, in part, to giving Newark students access to peers who reside outside of the district.

Missing Components of the Treatment

Resource Differences

As we note above, a true productivity analysis must account for differences in school inputs before coming to conclusions as to whether school outputs were realized due to greater efficiency. Yet the reports do not provide any context for understanding how Newark's schools vary in their access to resources.

This is particularly relevant to any discussion of the Newark charter sector. The reports correctly note that two large, national charter management organizations – KIPP/TEAM and Uncommon Schools – account for the majority of the enrollments in Newark charter schools, and for much of the growth in the sector. Several features, however, give these schools – and, in some cases, other Newark charters – important resource advantages.

It is notable that the current Superintendent of NPS, Christopher Cerf, publicly acknowledged, in 2015, the fiscal pressures charter schools put on his district:

> "Much of the budget pressure has come from payments the district is required to make to the city's charter schools. Cerf, a cheerleader for charter schools as state commissioner, yesterday acknowledged that some funding stop-gap is needed to help the district.
>
> "He said the mandatory funding for charter schools year to year is 'disproportionately hurting the district schools,' adding, 'We can't just turn the other way and let that happen.'"[81]

The fiscal advantages Newark's charter sector enjoys include:

Advantages in state aid allocation to charter schools: In fiscal year 2015-16, the Christie administration proposed a change to charter school funding that would essentially "hold harmless" state aid for charter schools, even as district state aid was kept flat.[82] According to the Education Law Center (ELC), Newark's charter schools received an additional $25 million in the final budget, even as NPS faced a $50 million budget gap.[83] Subsequent years have seen Newark's charters continue to enjoy this advantage in the allocation of state aid.[84]

Other governmental support: In 2013, the first phase of Teachers Village opened in downtown Newark. The mixed-use development, which would eventually become home to three of Newark's charter schools, benefitted from a reported $100 million in state and federal tax credits, as well as Qualified School Construction bonds and funding from other governmental sources.[85]

By way of contrast, in 2016 the public was informed that 30 NPS schools had unsafe levels of lead in the drinking water.[86] At the time, this was just the latest example of how the NPS physical plant, which includes many buildings the district itself has acknowledged were antiquated[87], had deteriorated to the point where many of the district's schools were unsafe.[88]

Newark charter schools receive other types of government support. KIPP NJ receives support through the federal AmeriCorps program, as does the Relay School of Education[89], which is closely aligned with both KIPP and Uncommon Schools.[90] Great Oaks Charter School also relies on AmeriCorps to provide tutors,[91] offering free housing as an incentive.[92]

Philanthropy: *The Prize* notes that over $57 million of the Zuckerberg donation and matching funds went specifically to support charter schools.[93] Yet that is only part of the total philanthropic contributions that annually flow to Newark's charter schools.

Just this past month, KIPP New Jersey announced it would collect between $7 and $9 mil-

lion from an anonymous donor on the sale of a single painting by artist Jean-Michel Basquiat.[94] While unusually large even by KIPP's standards, private donations are a standard source of revenue for the charter management organization (CMO). KIPP New Jersey's 2014 tax forms[95] show the organization collected nearly $1.7 million that year in donations and fundraising events. Friends of TEAM Academy, an aligned tax-exempt organization, collected over $1.8 million in donations in 2015, its latest available tax filing.

Uncommon Schools, the CMO of Newark's North Star Academy, reported over $12 million in contributions in its 2014 filing; it is unclear how much of that revenue was allocated towards North Star. Recent research suggests foundations that support charter school expansion have been directing more of their support toward CMOs[96]; the amounts collected by Uncommon and KIPP align with this finding.

The financial dealings of large CMOs like KIPP and Uncommon have become increasingly complex; a full accounting of how their activities may benefit their schools in Newark is beyond the scope of this review. But any meaningful productivity analysis of Newark's charter sector must include at least some acknowledgment of how philanthropy gives those schools a fiscal advantage over NPS.

Staff and wage differences: Our previous research on the New Jersey charter sector shows that charter schools employ staffs that are far less experienced than the staffs of their hosting school districts.[97] This creates an inherent fiscal advantage for the charter sector, as a less experienced teaching staff is inevitably a less expensive one.

One way a charter school may choose to use this advantage is to pay more competitive wages, relative to experience and credentials, so as to compensate teachers for longer school days and/or years. We further explore this advantage in Appendix C.

Curricular Narrowing

A substantial body of research has accumulated over the past several years that shows pressures brought on by high-stakes testing cause schools to narrow their curricula, focusing on tested subjects at the expense of instruction in non-tested domains.[98] In the absence of a comprehensive review of instructional programming in Newark's schools (and in the counterfactual schools used for comparison), it is impossible to state with certainty that gains in achievement growth in tested subjects for Newark can be attributed to this narrowing. Nonetheless, data does provide us with several interesting clues:

Instruction in domains other than math and English language arts requires teachers with a variety of certifications to provide that instruction. If school districts hire fewer teachers in different subjects proportionally, we can reasonably assume those schools are offering less expansive programming in those subjects. As an example: if we compare the "student loads" of art teachers at different schools, it gives us an indication of the relative depth and breadth of those schools' offerings in art.

We use staffing files to explore these differences in student loads between Newark's charter sector and NPS schools; see Appendix C for details. To summarize: relative to enrollment sizes, Newark's charter sector provides fewer personnel in the arts, physical education, world languages, social studies, and science than NPS schools. This suggests students who moved from district to charter schools also moved to schools with less expansive programs in non-tested subjects.

An additional clue as to whether Newark charter schools are narrowing their curriculum comes from the writings of the leaders of those schools. As Harvard education professor Daniel Koretz notes in his recent book, *The Testing Charade*[99], leaders of the CMOs that are running some of Newark's most prominent charter schools have made clear that gains in test scores are their top priority. In this passage, Koretz critiques the writings of Paul Bambrick-Santoyo[100] and Doug Lemov[101], who both hold leadership positions in the Uncommon Schools CMO, which manages Newark's North Star Academy:

> One of the rationales given to new teachers for focusing on score gains is that high-stakes tests serve a gatekeeping function, and therefore training kids to do well on tests opens doors for them. For example, in *Teaching as Leadership*[102] – a book distributed to many Teach for America trainees – Steven Farr argues that teaching kids to be successful on a high-stakes test "allows teachers to connect big goals to pathways of opportunity in their students' future." This theme is echoed by Paul Bambrick-Santoyo in *Leverage Leadership* and by Doug Lemov in *Teach Like a Champion*, both of which are widely read by new teachers. For example, in explaining why he used scores on state assessments to identify successful teachers, Lemov argued that student success as measured by state assessments is predictive not just of [students'] success in getting into college but of their succeeding there.
>
> Let's use Lemov's specific example to unpack this.
>
> To start, Lemov has his facts wrong: test scores predict success in college only modestly, and they have very little predictive power after one takes high school grades into account. Decades of studies have shown this to be true of college admissions tests, and a few more recent studies have shown that scores on states' high-stakes tests don't predict any better.
>
> However, the critical issue isn't Lemov's factual error; it's his fundamental misunderstanding of the link between better test scores and later success of any sort (other than simply taking another similar test). Whether raising test scores will improve students' later success – in contrast to their probability of admission – depends on *how* one raises scores. Raising scores by teaching well can increase students' later success. Having them memorize a couple of Pythagorian triples or the rule that *b* is the intercept in a linear equation[103] will increase their scores but won't help them a whit later.
>
> [...]

> Some of today's educators, however, make a virtue of this mistake. The[y] often tell new teachers that tests, rather than standards or a curriculum, *should* define what they teach. For example, Lemov argued that "if it's 'on the test,' it's also probably part of the school's curriculum or perhaps your state standards… It's just possible that the (also smart) people who put it there had a good rationale for putting it there." (Probably? Perhaps? Possible? Shouldn't they look?) Bambrick-Santoyo was more direct: "Standards are meaningless until you define how to assess them." And "instead of standards defining the sort of assessments used, the assessments used define the standard that will be reached." And again: "Assessments are not the end of the teaching and learning process; they're the starting point."
>
> They are advising new teachers to put the cart before the horse."[104]

"Teaching to the Test"

Koretz's critique begs a question: just how much of the effect found in the reports' research can be attributed to true gains in learning due to instructional and curricular improvements, and how much is due to simply "teaching to the test"?

The reports' clearly show that the largest gains in test score growth were found between 2014 and 2015, when New Jersey shifted from the NJASK to the PARCC. Recent research suggests that changes in assessments do not generally affect math value-added, but in some cases do affect ELA.[105] This aligns with the findings of the reports: value-added only increased in ELA, and not in math.

The reports make a notable observation about the shift and district-level VAM outcomes:

> 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. However, 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, this implies 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.**[106]

The last sentence implies a remarkable assumption: that the difference between Newark's relative standing in value-added between 2014 and 2015 is due to the PARCC being a "better"

measure of instructional and curricular efficacy than the NJASK. The reports cite research by the authors that purports to show the PARCC is more "instructionally sensitive" than previous state accountability tests. This research does not include New Jersey, but even if it did: it does not provide any evidence that achievement gains on the PARCC, relative to other statewide tests, demonstrate "better" teaching that leads to "better" non-test outcomes.

We readily concede such a claim would be beyond the scope of these reports. But when most of the relative gains in value-added are found in the same year New Jersey switched its tests, great caution is warranted before drawing any conclusions about the causal effects of programs on true student achievement gains. We explore this issue further in Appendix B.

Effect Sizes

The reports find no statistically significant change in value-added for math, which immediately calls into question the efficacy of the Newark-Zuckerberg reforms. The reports do find an overall gain in 2016 (relative to 2010/2011) in value-added of 0.070 standard deviations in ELA.[107] As we note in Appendix B, however, this gain is relative to the entire state. Comparing achievement growth in Newark to distant schools is highly problematic for a number of reasons. When we instead compare Newark's achievement growth to other economically disadvantaged districts also located in Essex County (using the New Jersey Department of Education's (NJDOE) preferred growth measures), we find that Newark's schools, both charter and district, experienced similar gains. This similarity suggests Newark's small gain in ELA relative to the rest of the state is suspect; it likely is due regional differences, and not policies implemented solely in Newark.

To make the case this gain is practically significant, the reports offer this comparison:

> Between the baseline period (2010/2011) and the most recent available year (2016), average value-added in Newark grew by 0.07 SD in English. In English, that is a sizeable gain, comparable to the impact of being assigned to an experienced versus novice teacher.[108]

The only citation offered to validate this comparison is two of the authors' own work: a policy brief published over ten years ago.[109] From that brief: "The evidence suggests that the average 'value-added' of novices is about 4 percentile points lower than for teachers with two years of experience."[110]

This research is based solely on a dataset of Los Angeles teachers, and only compares teachers with two years of experience to brand-new teachers. This is a rather narrow piece of evidence to make the reports' more sweeping comparison. A more comprehensive review of the literature on teacher experience from 2016[111] finds that teacher experience beyond the first few years can substantially impact student achievement. Depending on the study, its dataset, and its methods, the effect sizes can be in excess of 0.07 SD.

Our point here is to note that very little context is given in the reports for the lay reader to

make a considered judgment as to whether the VAMs employed find a *practically* significant effect. In a very basic interpretation, for example, 0.07 SD moves a comparison group from the 50th to the 53rd percentile.[112] We believe lay readers would see this interpretation as far less impressive than the limited context offered by the reports.

Consider the following hyperbolic quotes from the announcement of the donation on *The Oprah Winfrey Show*, September 24, 2010[113]:

> Cory Booker, at the time Mayor of Newark: "Well, we've been talking for quite some time about creating **a bold, new paradigm for educational excellence in the country**, to show the way."
>
> Chris Christie, Governor of New Jersey: "… to develop this entirely new plan of how to reform the education system in Newark and **create a national model.**"
>
> Mark Zuckerberg, Founder and CEO of Facebook: "We're setting up a $100 million challenge grant so that Mayor Booker and Governor Christie can have the flexibility that they need to implement new programs in Newark and really make a difference, and **turn Newark into a symbol of educational excellence for the whole nation**."

No matter how one chooses to interpret the sizes of the effects found in this study, one would be hard-pressed to suggest a 0.07 SD effect in ELA and no effect in math are "educationally meaningful" in the context of the goals stated above.

The Larger Context of Newark School Reform

In Appendix A, we provide a more comprehensive historical view of Newark school policy reform than is provided in the reports. We note that the Abbott rulings, the expansion of charter schools, and the 2008 School Funding Reform Act have consequences that even today influence student achievement in the city.

In general, while Newark and other urban districts in New Jersey saw substantial increases in funding during the "scaling up" period of the Abbott reforms, Newark also experienced changes in resource allocation and student characteristics that were substantially different from surrounding public school districts. These changes must be accounted for in a true productivity analysis of the later Newark "reforms."

But even if we were to include these and many other factors related to school inputs and student characteristics into the reports' models, and even if we could fully describe and isolate the reforms attributable to the Zuckerberg donation, we would still not fully capture their effects on Newark students and families. This is because test scores are the only outputs used in the reports' models.

We don't dispute the importance of test scores in conducting an analysis; however, we do note there are many other factors that must be considered when evaluating the overall effect

of the Newark "reforms" of the last five years. Consider:

- According to the U.S. Department of Education Office of Civil Rights, a preliminary investigation of the 2012 school closings "…found that a 'significantly disproportionate' number of black students were affected compared to their white peers."[114] The district was compelled to sign an agreement that required the district to "…identify any students harmed by the reorganization, and take steps to remedy the adverse effects."[115]

- In 2012, the Education Law Center settled a lawsuit on behalf of six Newark students that called for improvements in access to special education services.[116] However, according to a 2015 report from the Education Law Center[117], at least 350 Newark students "…did not receive timely evaluations for special education over the last six months…"

- In April of 2014, 77 members of Newark's clergy released a position statement on One Newark, the universal enrollment system, calling for a moratorium on its implementation.[118] "The One Newark Public School Plan, as currently proposed, is already producing irreversible changes and fomenting widespread outrage. It has caused unnecessary instability, in the Newark public school system, as well as the lives of thousands of its families. The disruptive and divisive nature of the One Newark Public School Plan could have catastrophic and far-reaching consequences for the children of Newark, the reputation of the State of New Jersey, and have implications for urban education nationally."

- In February of 2014, then state superintendent Cami Anderson released a letter stating she would no longer attend meetings of the Newark School Advisory Board, claiming "dysfunction" had made the meetings "a bad example for our children."[119] Anderson eventually resigned in 2015.[120]

- In the spring of 2015, eight Newark high school students occupied NPS central offices for four days, demanding Anderson's resignation.[121] That May, thousands of Newark students participated in a mass demonstration, walking out of classes and blocking city traffic.[122]

- In 2015, the USDOE found NPS had violated portions of New Jersey's waiver from the No Child Left Behind Act, failing to exercise proper oversight of its improvement plans.[123]

- While the Newark Teachers Union and NPS did eventually reach a new contract settlement in 2017[124], negotiations were acrimonious.[125] Tensions between State Superintendent Chris Cerf and the NTU reportedly remain high today.[126]

Of course, none of these factors can necessarily be traced back to the Zuckerberg donation. But there is little doubt that during the time of the reports' research, NPS has been characterized by strife and hostility. We believe no analysis of the city's schools during this time can be complete without some acknowledgement of this discord.

VI. Review of the Validity of the Findings and Conclusions

In any fair reading of the reports' conclusions, the Newark "reforms" of recent years, which correctly or incorrectly credit the Zuckerberg donation, may have produced, at best, small gains in English language arts test outcomes, and no gains in math.

As our review of the methods and exploration of other data shows, however, even this small gain in ELA is suspect. Too many factors – particularly, the resource and student characteristic differences between Newark's district and charter schools – are simply not accounted for in the models presented.

In addition, there is sufficient reason, in our view, to believe other factors not fully accounted for by the reports – historical trends, curricular narrowing, measurement artifacts due to changing assessments, and others – make the VAM estimations and their decomposition suspect.

VII. Usefulness of the Report for Guidance of Policy and Practice

We are long-time observers and researchers of New Jersey education policy (see Appendix D). As such, we can state one thing for certain: the story of Newark education reform is highly complex. While we regularly rely on quantitative methods in our own research, we readily acknowledge its limitations, as we have seen time and again that important information is omitted in such methods. Crude variables for student characteristics, inadequate measures of school resources, and error-prone test outcomes limited to two domains of learning are the realities we deal with regularly in our work.

Given these limits, we are already skeptical when we approach econometric research that makes causal claims about the effects of particular policy interventions on student achievement. That skepticism grows when evaluating research that does not clearly define the treatment in question, omits important factors we know affect student learning, and finds what can, at best, be described as small effect sizes.

The reports conclude: "The experience in Newark has shown that re-allocation of market share can be an important contributor to productivity growth in K-12 education, as it has been in many other industries."[127] We respectfully disagree, for the following reasons:

- The reports are not productivity research. They do not account for resource differences between schools and school districts. Because these differences and other important factors, such as unmeasured student characteristic differences, are omitted from the models used, there is ample reason to believe the effects estimated are biased.
- Even if we accept the validity of the estimates, the effect sizes found are practi-

cally non-significant, especially when judged against the claims of those who attempted to initiate the recent reforms through the Zuckerberg donation.

- The effects of the reforms in question cannot be properly evaluated without accounting for the disruption suffered by the citizens and students of Newark.

For these reasons, we believe these reports do not provide useful guidance for policy makers in Newark, other communities in New Jersey, or the rest of the nation.

Endnotes

1 For example, see:
Barnum, M. (2017, October 16). The $100 million question: Did Newark's school reforms work? New study finds big declines, then progress. *Chalkbeat*. Retrieved November 26, 2017, from https://www.chalkbeat.org/posts/us/2017/10/16/the-100-million-question-did-newarks-school-reforms-work-new-study-finds-big-declines-then-progress/

Toppo, G. (2017, October 16). What did Zuckerberg's $100 million buy in Newark? A bit of progress. *USA Today*. Retrieved November 26, 2017, from https://www.usatoday.com/story/news/2017/10/16/what-did-zuckerbergs-100-million-buy-newark-bit-progress/769536001/

2 Russakoff, D. (2016). *The Prize; Who's in charter of America's schools?* New York, NY: Houghton, Mifflin, Harcourt. p.35.

3 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.5.

4 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.3.

5 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.7.

6 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 26.

7 This drop is not statistically significant; see Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 44, Table 2.

8 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 26.

9 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 22: "*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.15 However, 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.*"

10 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *Assessing the Impact of the Newark Education Reforms: The Role of Within-School Improvement vs. Between-School Shifts in Enrollment*. Center for Education Policy Research, Harvard University. Retrieved November 26, 2017, from https://cepr.harvard.edu/evaluating-newark-school-reform, p. 4.

11 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.29.

12 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.

13 Chetty, R., John N. Friedman, J. N., & Jonah E. Rockoff, J. E. (2014). Measuring the impacts of teachers I: evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9): 2593-2632.

14 If not unique to Newark or the time period in question, they must at least be measured relative to the rest of the study population and study years.

15 Grade 3 is the first year that students take statewide assessments in New Jersey; that grade is not evaluated on growth because it is the baseline.

16 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 28

17 "*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; changes in the treatment of incomplete items in the scoring of NJASK and PARCC. None of these hypotheses can account for the rise in achievement by many low-income districts in New Jersey."* Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 28

18 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 28. "Abbott" districts are the original 31 low-income districts which were plaintiffs in a series of cases which led to school funding reforms in New Jersey, culminating in the 2008 School Funding Reform Act. See: Education Law Center (n.d.) The history of Abbott v. Burke. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/cases/abbott-v-burke/abbott-history.html

19 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.26.

20 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 7.

21 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 5.

22 Center for Research on Education Outcomes (2015). Urban charter school study report on 41 regions. Retrieved November 26, 2017, from https://urbancharters.stanford.edu/download/Urban%20Charter%20School%20Study%20Report%20on%2041%20Regions.pdf

23 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark:*

Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 19.

24 Chandra, A., Finkelstein, A., Sacarny, A., & Syverson, C. (2016). Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector. *American Economic Review*, *106*(8), 2110–2144. https://doi.org/10.1257/aer.20151080

25 Baily, M.N., Hulten, C., Campbell, D., Bresnahan, T., & Caves, R.E. (1992). Productivity Dynamics in Manufacturing Plants. *Brookings Papers on Economic Activity. Microeconomics*, *1992*, 187–267. https://doi.org/10.2307/2534764

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, *98*(1), 394–425. https://doi.org/10.1257/aer.98.1.394

26 Foster, L., Haltiwanger, J.C., & Krizan, C.J. (2001). Aggregate Productivity Growth. Lessons from Microeconomic Evidence. In *New Developments in Productivity Analysis*. Retrieved November 26, 2017, from http://www.nber.org/chapters/c10129

27 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 4.

28 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 5.

29 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 15.

30 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.

31 Field, A.J. (2008) Productivity. *The concise encyclopedia of economics, 2nd edition*. Retrieved November 26, 2017, from http://www.econlib.org/library/Enc/Productivity.html "*The growth of productivity—output per unit of input—is the fundamental determinant of the growth of a country's material standard of living.*" For a short introduction to "productivity" in layman's terms, see this video from the US Bureau of Labor Statistics: https://www.bls.gov/bls/productivity.htm

32 Burkhead, J., & Hennigan, P.J. (1978). Productivity Analysis: A Search for Definition and Order. *Public Administration Review*, *38*(1), 34–40. https://doi.org/10.2307/975408

33 Rice, J.K. & Schwartz, A.E. (2015). Toward and understanding of productivity in education. In Ladd, H.F. & Goertz, M.E. (Eds.), *Handbook of research in education finance and policy, 2nd edition* (125-140). New York, NY: Routledge.

34 For a comprehensive review of this research, see: Baker, B.D. (2016). *Does Money Matter in Education? Second Edition*. Washington, D.C.: Albert Shanker Institute. Retrieved November 26, 2017, from http://www.shankerinstitute.org/resource/does-money-matter-second-edition
We note that since the publication of this review, additional studies have confirmed the influence of school funding policy on student outcomes, including:
Lafortune, J.R., Rothstein, J., & Schanzenbach, D.W. (2016). *School Finance Reform and the Distribution of Student Achievement*. Retrieved November 26, 2017, from http://irle.berkeley.edu/workingpapers/100-16.pdf

Baker, B.D., & Weber, M. (2016). Beyond the Echo-Chamber: State Investments and Student Outcomes in US Elementary and Secondary Education. *Journal of Education Finance*, *42*(1), 1–27.

35 Baker, B.D. (2014). America's Most Financially Disadvantaged School Districts and How They Got That Way: How State and Local Governance Causes School Funding Disparities. Washington, DC: Center for American Progress. Retrieved November 26, 2017, from https://cdn.americanprogress.org/wp-content/uploads/2014/07/BakerSchoolDistricts.pdf

36 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.

37 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). School District Reform in Newark: Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922*)*. National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 8.

38 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). *Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research. Retrieved November 26, 2017, from https://cdn.americanprogress.org/wp-content/uploads/2014/07/BakerSchoolDistricts.pdf http://www.air.org/resource/newark-public-schools-and-newark-teachers-union-teacher-contract-evaluation

39 https://www.aft.org/press-release/newark-teachers-union-wins-huge-arbitration-award - This webpage includes a link to the text of the arbitrator's September 13, 2017 ruling on NTU's grievance.

40 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). School District Reform in Newark: Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 34.

41 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). School District Reform in Newark: Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 7.

42 Russakoff, D. (2016). The Prize; Who's in charter of America's schools? New York, NY: Houghton, Mifflin, Harcourt. p. 121.

43 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). School District Reform in Newark: Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.8-9.

44 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). School District Reform in Newark: Within- and Between-School Changes in Achievement Growth (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.

45 Mooney, J. (2012, October 19). Newark teachers contract, a game-changer for NJ's often-troubled school district. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/12/10/19/newark-teachers-contract-a-game-changer-for-nj-s-often-troubled-school-district/

46 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). *Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research. Retrieved November 26, 2017, from https://cdn.americanprogress.org/wp-content/uploads/2014/07/BakerSchoolDistricts.pdf http://www.air.org/resource/newark-public-schools-and-newark-teachers-union-teacher-contract-evaluation

47 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.8.

48 American Federation of Teachers (2017, October 3). Press Release: Newark Teachers Union Wins Huge Arbitration Award. Retrieved November 26, 2017, from

https://www.aft.org/press-release/newark-teachers-union-wins-huge-arbitration-award

49 Tomassini, J. (2012, Oct. 18) Facebook Donation to Pay for Newark Teacher Performance Bonuses. *EdWeek Market Brief*. Retrieved November 26, 2017, from https://marketbrief.edweek.org/marketplace-k-12/newark_will_fund_part_of_new_teacher_contract_with_facebook_donation/

50 Mooney, J. (2012, October 19). Facebook fund helps seal the deal for Newark teachers' contract. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/12/11/16/facebook-fund-helps-seal-the-deal-for-newark-teachers-contract/

For an accounting of the ambiguity in these initial news stories, see: Weber, M. (2013, August 28). Updated: Newark's teachers get screwed on merit pay [Blog post]. Retrieved November 26, 2017, from http://jerseyjazzman.blogspot.com/2013/08/newarks-teachers-get-screwed-on-merit.html

51 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.7.

52 Memorandum of agreement between the Newark state operated school district and Newark Teachers Union, October 18, 2012. Retrieved November 26, 2017, from http://assets.njspotlight.com/assets/12/1025/0021

53 Fleisher, L. (2013, August 29). Newark's merit-pay plan begins. *The Wall Street Journal*. Retrieved November 26, 2017, from https://www.wsj.com/articles/newarks-meritpay-plan-begins-1377481599?tesla=y

54 Mulvihill, G. (2015, October 1) Newark teachers union not in love with breakthrough contract. *Chicago Tribune*. Retrieved November 26, 2017, from http://www.chicagotribune.com/news/sns-bc-us--newark-schools-facebook-money-20151001-story.html

55 Taylor, D. (2014, Jan 21) Union says Newark principals suspended because they opposed district changes. *NJTV News*. Retrieved November 26, 2017, from https://www.njtvonline.org/news/video/union-says-newark-principals-suspended-because-they-opposed-district-changes/

56 Meyer, P. (2013). Newark's Superintendent Rolls Up Her Sleeves and Gets to Work. *EducationNext, 13*(1). Retrieved November 26, 2017, from http://educationnext.org/newark's-superintendent-rolls-up-her-sleeves-and-gets-to-work/

57 Russakoff, D. (2016). *The Prize; Who's in charter of America's schools?* New York, NY: Houghton, Mifflin, Harcourt. p. 121.

58 Russakoff, D. (2016). *The Prize; Who's in charter of America's schools?* New York, NY: Houghton, Mifflin, Harcourt. p. 223. Russakoff notes that some of the funds committed were used "on rewards to individual principals"; no further details are offered.

59 Calefati, J. (2012). Principals selected for 8 struggling schools in Newark. *The Star Ledger*. Retrieved November 26, 2017, from http://www.nj.com/news/index.ssf/2012/05/principals_selected_for_8_stru.html

McGlone, P. (2013). Newark school restructuring includes plans to put charters in district buildings. *NJ.com*. Retrieved November 26, 2017, from http://www.nj.com/education/2013/12/newark_school_plan_restructure.html

60 Kamenetz, A. (2013, April 15). Cami Anderson and Zuckerberg, the unlikely duo behind Newark schools' revitalization. *Fast Company*. Retrieved November 26, 2017, from https://www.fastcompany.com/3007765/cami-anderson-and-zuckerberg-unlikely-duo-behind-newark-schools-revitalization

61 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). *Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research. Retrieved November 26, 2017, from

https://cdn.americanprogress.org/wp-content/uploads/2014/07/BakerSchoolDistricts.pdf http://www.air.org/resource/newark-public-schools-and-newark-teachers-union-teacher-contract-evaluation, p. 11.

62 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report. American Institutes for Research. Retrieved November 26, 2017, from https://cdn.americanprogress.org/wp-content/uploads/2014/07/BakerSchoolDistricts.pdf http://www.air.org/resource/newark-public-schools-and-newark-teachers-union-teacher-contract-evaluation, p. A-7.

63 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p.24.

64 Mooney, J. (2010, June 17). NJ Board of Education adopts Common Core standards. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/10/0616/2136/

65 Clark, A. (2016, May 5). N.J. revises, renames Common Core academic standards. *The Star-Ledger*. Retrieved November 26, 2017, from http://www.nj.com/education/2016/05/nj_common_core_standards_christie.html

66 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 8, 21.

67 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 8.

68 https://www.edreports.org/ela/expeditionary-learning/index.html

69 Kane, T.J., Owens, A.M., Marinell, W.H., Thal, D.R.C., & Staiger, D.O. (2016). *Teaching Higher: Educators' perspectives on common core implementation*. Center for Education Policy Research, Harvard University.

70 The states are: Delaware, Maryland, Massachusetts, New Mexico, and Nevada.

71 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 27.

72 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 27.

73 Weber, M. & Baker, B.D. (2014). *An Empirical Critique of "One Newark."* New Brunswick, NJ: New Jersey Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/01/24/new-report-an-empirical-critique-of-one-newark/

74 Weber, M. (2014). *Buyer Beware: One Newark and the Market For Lemons*. New Brunswick, NJ: New Jersey Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/05/23/buyer-beware-one-newark-and-the-market-for-lemons/

75 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.

76 Center for Research on Education Outcomes (2015). *Urban charter school study report on 41 regions*. Retrieved November 26, 2017, from https://urbancharters.stanford.edu/download/Urban%20Charter%20School%20Study%20Report%20on%2041%20Regions.pdf

77 Baker, B.D. (2012) The Secrets to Charter School Success in Newark: Comments on the NJ CREDO Report [Blog post]. Retrieved November 26, 2017, from https://schoolfinance101.wordpress.com/2012/11/27/the-secrets-to-charter-school-success-in-newark-comments-on-the-nj-credo-report/

78 For an excellent review of the literature on this topic, see: Altenhofen, S., Berends, M., & White, T.G. (2016). School Choice Decision Making Among Suburban, High-Income Parents. *AERA Open*, *2*(1), 1–14. https://doi.org/10.1177/2332858415624098

79 Weber, M. (2015). *One Newark: Choosing "Great" Schools, or Merely Segregated Ones?* New Brunswick, NJ: New Jersey Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2015/04/24/one-newark-choosing-great-schools-or-merely-segregated-ones/

80 Weber, M.A. (2014). *Buyer Beware: One Newark and the Market For Lemons*. New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/05/23/buyer-beware-one-newark-and-the- market-for-lemons/

81 Mooney, J. (2015, Oct. 8). Board hears update on status of two state-run school districts. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/15/10/08/board-hears-update-on-status-of-two-state-run-school-districts/

82 Education Law Center (2015). Governor Christie wants to hold charters harmless, but not district schools. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/nj-charter-schools/governor-christie-wants-to-hold-charters-harmless-but-not-district-schools.html

83 Education Law Center (2015). Final NJ budget: Charter and private schools get more, no increase for district schools. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/school-funding/final-nj-budget-charter-and-private-schools-get-more-no-increase-for-district-schools.html

84 Education Law Center (2016). Governor Christie's FY17 school aid budget is wholly inadequate again. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/school-funding/governor-christies-fy17-school-aid-budget-is-wholly-inadequate-again.html

85 De Poto, T. (2013, September 25). Booker, Christie cut ribbon to officially open Teachers Village in Newark. *The Star-Ledger*. Retrieved November 26, 2017 from http://www.nj.com/business/index.ssf/2013/09/newark_christie_booker_teacher.html

86 Braun, B. (2016, March 9). Newark school children were drinking lead contaminated water–but for how long? [Blog post]. Retrieved November 26, 2017, from http://www.bobbraunsledger.com/newark-school-children-were-drinking-lead-contaminated-water-but-for-how-long/

87 Newark Public Schools (2013, December). *Building a system: One Newark*.

88 Education Law Center (2014). Top 20 worst Newark school buildings. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/school-facilities/top-20-worst-newark-school-buildings.html

89 State of New Jersey, Department of State (n.d.). New Jersey AmeriCorps 2016-2017 programs. Retrieved November 12, 2017, from http://www.nj.gov/state/programs/dos_program_americorps_programs.html

90 https://www.relay.edu/news/2014/27/11/freakonomics-features-relay-podcast

91 https://my.americorps.gov/mp/listing/viewListing.do?id=63938

92 https://greatoakscharter.org/join/great-oaks-tutor-corps/about-tutor-corps/

93 Russakoff, D. (2016). *The Prize; Who's in charter of America's schools?* New York, NY: Houghton, Mifflin, Harcourt, p. 224.

94 Freeman, N. (2017, Oct. 12). Basquiat's 'Red Skull,' sold at Christie's in London for $21.5 M, will fund new charter schools in New Jersey and Miami. *Art News.* Retrieved November 26, 2017, from http://www.artnews.com/2017/10/12/basquiats-red-skull-sold-at-christies-in-london-for-21-5-m-will-fund-new-charter-schools-in-new-jersey-and-miami/

95 KIPP NJ, Friends of TEAM, and Uncommon tax forms are available at: http://www.guidestar.org

96 Ferrare, J.J., & Setari, R.R. (2017). Converging on Choice: The Interstate Flow of Foundation Dollars to Charter School Organizations. *Educational Researcher*, *XX*(X), 1–12. https://doi.org/10.3102/0013189X17736524

97 Weber, M.A. (2016). *New Jersey Charter Schools: A Data-Driven View, Part II - Finances and Staffing*. Princeton, NJ: Daniel Tanner Foundation. Retrieved November 26, 2017, from http://www.saveourschoolsnj.org/nj-charter-school-data/

98 Dee, T.S., Jacob, B., & Schwartz, N.L. (2013). The Effects of NCLB on School Resources and Practices. *Educational Evaluation and Policy Analysis*, *35*(2), 252–279. https://doi.org/10.3102/0162373712467080

Murnane, R.J., & Papay, J.P. (2010). Teachers' Views on No Child Left Behind: Support for the Principles, Concerns about the Practices. *Journal of Economic Perspectives*, *24*(3), 151–166. https://doi.org/10.1257/jep.24.3.151

99 Koretz, D. (2017) *The testing charade; Pretending to make schools better*. Chicago, IL: University of Chicago Press.

100 http://www.uncommonschools.org/bio/1017/paul-bambrick-santoyo

101 http://www.uncommonschools.org/our-approach/thought-leadership/teach-like-a-champion-book-doug-lemov

102 Farr, S. (2010). *Teaching as leadership; The highly effective teacher's guide to closing the achievement gap*. San Francisco: Josey-Bass. We note here that Russakoff reports that Teach for America received $1 million of the Zuckerberg donation "to train teachers for positions in Newark district and charter schools." (Russakoff, D. (2016). *The Prize; Who's in charter of America's schools?* New York, NY: Houghton, Mifflin, Harcourt. p. 224)

103 A "Pythagorean Triple" is a memorized ratio that conforms to the Pythagorean theorem regarding the ratio of the sides of a right triangle. Koretz critiques the linear intercept rule, noting that *b* is often taught as the intercept of an equation in high school, but is usually the coefficient of an equation in college courses. In both cases, Kortez contends test prep strategies keep students from gaining a full understanding of the concepts being taught. See: Koretz, D. (2017) *The testing charade; Pretending to make schools better*. Chicago, IL: University of Chicago Press. pp. 104-108.

104 Koretz, D. (2017) *The testing charade; Pretending to make schools better*. Chicago, IL: University of Chicago Press. p. 114-115.

105 Backes, B., Cowan, J., Goldhaber, D., Koedel, C., Miller, L.C., & Xu, Z. (2017). The Common Core Conundrum: To What Extent Should We Worry That Changes to Assessments Will Affect Test-Based Measures of Teacher Performance? *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2017.10.004

106 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 22

107 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 43. Statistically significant at the $p<0.01$ level.

108 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 26.

109 Gordon, R., Kane, T.J., & Staiger, D.O. (2006). *Identifying Effective Teachers Using Performance on the Job*. Brookings Institute. Retrieved November 26, 2017, from https://www.brookings.edu/research/identifying-effective-teachers-using-performance-on-the-job/

110 Gordon, R., Kane, T.J., & Staiger, D.O. (2006). *Identifying Effective Teachers Using Performance on the Job*. Brookings Institute. Retrieved November 26, 2017 from https://www.brookings.edu/research/identifying-effective-teachers-using-performance-on-the-job/, p. 14.

111 Kini, T., & Podolsky, A. (2016). *Does Teaching Experience Increase Teacher Effectiveness? A Review of the Research*. Retrieved November 26, 2017, from https://learningpolicyinstitute.org/product/does-teaching-experience-increase-teacher-effectiveness-review-research. Pages 15 through 18 summarize several studies that show effect sizes on returns to experience ranging from 0.08 to 0.21 SD.

112 This is the most generous interpretation possible based on changes in percentiles; however, the change is actually less because Newark students did not start at the 50th percentile in the baseline year.

113 Mark Zuckerberg Announces $100 Million Grant [video] (2010). Retrieved November 26, 2017, from http://www.oprah.com/own-oprahshow/mark-zuckerbergs-big-announcement-video. The transcript is our own.

114 Ivers, D. (2015, December 16). Newark schools, feds strike deal to halt probe into civil rights complaints. *NJ Advance Media for NJ.com*. Retrieved November 26, 2017, from http://www.nj.com/essex/index.ssf/2015/12/newark_schools_strike_deal_with_feds_to_halt_probe.html

115 Ivers, D. (2015, December 16). Newark schools, feds strike deal to halt probe into civil rights complaints. *NJ Advance Media for NJ.com*. Retrieved November 26, 2017, from http://www.nj.com/essex/index.ssf/2015/12/newark_schools_strike_deal_with_feds_to_halt_probe.html

116 Education Law Center (2012). Education Law Center, Gibbons P.C., Newark Public Schools and State of NJ settle decade-long class action. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/special-education/education-law-center-gibbons-p.c.-newark-public-schools-and-state-of-nj-settle-decade-long-class-action.html

117 Education Law Center (2015). Hundreds of Newark children still not receiving special education services on time. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/news/archives/special-education/hundreds-of-newark-children-still-not-receiving-special-education-services-on-time.html

118 Mooney, J. (2014, April 22). Plea by Newark clergy ratchets up debate over school reforms. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/14/04/21/plea-by-newark-clergy-ratchets-up-debate-over-school-reforms/

119 Mooney, J. (2014, April 22). Plea by Newark clergy ratchets up debate over school reforms. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/14/02/26/fine-print-cami-anderson-open-letter-to-newark-families-says-she-ll-no-longer-attend-school-board-meetings/

120 Zernike, K. (2015, June 22). Cami Anderson, Picked by Christie, is out as Newark Schools Superintendent. *The New York Times*. Retrieved November 26, 2017, from https://www.nytimes.com/2015/06/23/nyregion/newark-schools-superintendent-is-stepping-down.html?_r=0

121 Layton, l. (2015, March 3). Chris Christie's bold plan to remake public schools is running into trouble. *The Washington Post*. Retrieved November 26, 2017, from

https://www.washingtonpost.com/local/education/newark-officials-seek-obamas-help-to-stop-chris-christies-school-plan/2015/03/03/c4e23dc0-b7b2-11e4-9423-f3d0a1ec335c_story.html?tid=a_inl&utm_term=.e8a9335246a5

122 Wichert, B. (2015, May 22). Thousands of Newark students leave school in protest, block major intersection. *NJ Advance Media for NJ.com*. Retrieved November 26, 2017, from http://www.nj.com/essex/index.ssf/2015/05/newark_students_leave_school_in_protest_block_majo.html

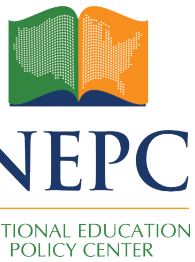
123 Ivers, D. (2015, June 29). Cerf, Anderson may have violated education agreement in failing Newark schools, feds say. *Advance Media for NJ.com*. Retrieved November 26, 2017, from http://www.nj.com/essex/index.ssf/2015/06/cerf_anderson_violated_education_agreement_in_fail.html

124 Mooney, J. (2017, May 16). Newark Teachers Union inks new contract – without Oprah, Zuckerberg. *NJ Spotlight*. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/17/05/16/newark-teachers-union-inks-new-contract-without-oprah-zuckerberg/

125 Bonamo, M. (2017, January 18). Newark superintendent calls out teachers' union boss over contract stalemate. *TAPinto.net*. Retrieved November 26, 2017, from https://www.tapinto.net/articles/newark-superintendent-calls-out-teachers-union-b

126 Braun, B. (2017, October 18). Newark schools chief to teacher union head: Don't need you [blog post]. Retrieved November 26, 2017, from http://www.bobbraunsledger.com/newark-schools-chief-to-teacher-union-head-dont-need-you/

127 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 29.



Appendices for NEPC Review: School District Reform in Newark (National Bureau of Economic Research, October 2017) and Impact of the Newark Education Reforms (Center for Education Policy Research, Harvard University, September 2017)

Reviewers:

Mark Weber and Bruce D. Baker
Rutgers, the State University of New Jersey

December 2017

Appendix A
The History and Context of Newark Education Reforms

In this appendix we address the significant omissions to the reports' timeline of reforms and interventions affecting Newark Public Schools and charter schools operating in Newark. The NBER paper indicates[1] that the major events affecting schooling in Newark included:

- 1985: School finance litigation, resulting in court ordered reform presumably affecting funding for Newark (not measured, cited or reported by authors).
- 1995: The state government takes over Newark Public Schools.
- 2010: Mark Zuckerberg appears on the Oprah show to announce his reform-focused gift to Newark.

Missing from this timeline, however, are several substantive reforms that occurred statewide, with specific effects on Newark's schools, between 1998 and the time of the Zuckerberg donation. Those reforms fall into three major categories, which set the stage for more recent reforms. Further, these reforms and their consequences provide relevant context for understanding counterfactuals in any analysis of Newark Public Schools. Those reforms include:

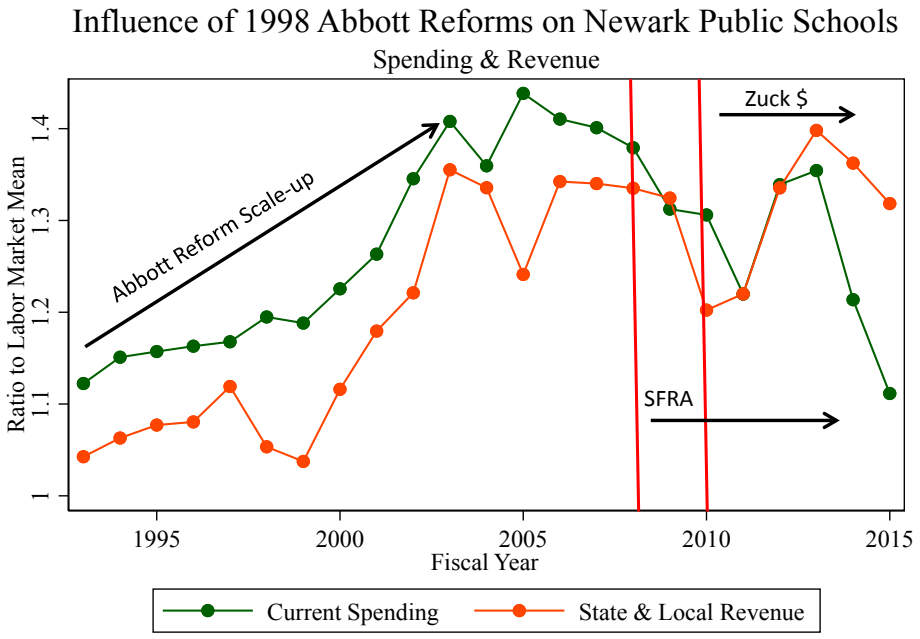
1. 1998 Abbott rulings, which led to a substantial infusion of funding into Newark and other Abbott districts, including the introduction and expansion of state-funded universal pre-k programs, a large infusion of funding for capital infrastructure, and substantial infusion of general operating state aid;[2]
2. The introduction and expansion of charter schooling, and opening of what have become the state's and city's most well-heeled operators, including TEAM/KIPP Academy, North Star/Uncommon Academy, and Robert Treat Academy;[3]
3. The 2008 adoption of the School Funding Reform Act (SFRA),which initiated the scaling back of funding to high-poverty urban districts including Newark, and, until the recession a few years later, began distributing more aid to non-Abbott, high-poverty districts.[4]

The recession, which followed these reforms and had effects continuing through the Zuckerberg gift era, also had substantive effects on school resources in Newark and other districts.[5] In some cases these effects fell unevenly across districts, including the court-imposed reinstatement of five percent aid cuts to Abbott districts, but not to non-Abbott[6] (including many high poverty) districts in 2010.

Listing events that presumably influenced resources is less informative than actually measuring changes in resources, so here we do the latter. Figure A1 shows Newark Public Schools funding levels as a ratio to their labor market average (metro area surrounding Newark) during the scale up and eventual pull back of Abbott-1998 reforms, including adoption of SFRA and the recession. Revenue and spending went from approximately average to 30% to 40% above average between 1995 and 2005, then leveled, then plummeted back to 20%

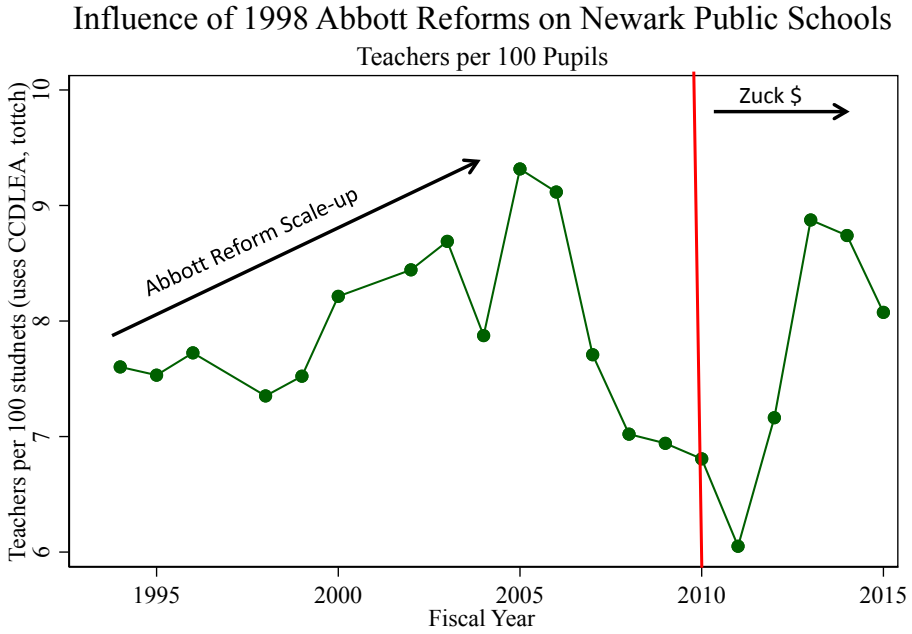
above average during the recession. The initial infusion of funding was coupled with an infusion of staff, as seen in Figure A2. But staffing reductions followed funding reductions during the recession.

Figure A1



Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System*. Retrieved November 26, 2017, from: http://www.schoolfundingfairness.org/data-download

Figure A2



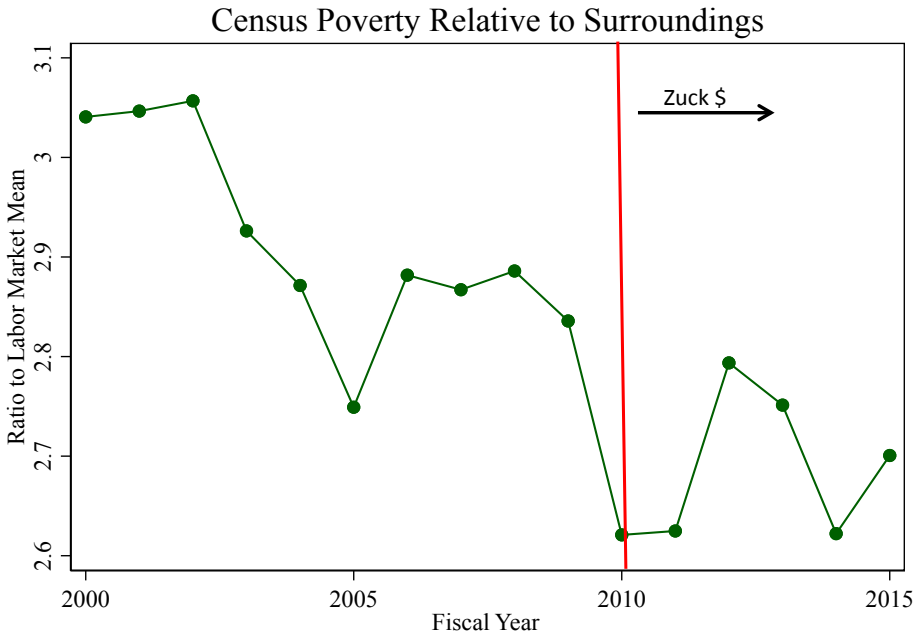
Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System*. Retrieved November 26, 2017, from: http://www.schoolfundingfairness.org/data-download

During much of this period, charter school enrollments were gradually on the rise.[7] Charter school enrollments escalated during the same time period in which NPS faced increased fiscal stress due to aid cuts and freezes during the recession. During this same period, charter school funding was held harmless, and Newark's charter sector grew to be the largest (as a share of enrollment) of any district statewide.[8]

Thus, when considering any comparisons across Newark district schools, Newark charter schools, schools in other Abbott districts and other schools in New Jersey, one must consider the significantly different treatments across these groups throughout this period. Newark district schools faced cuts and freezes while charters were buffered. Other Abbott districts benefited from the restoration of 2010 cuts, while poor non-Abbotts did not. While NPS also benefited from the restoration, NPS was experiencing much greater transfer of resources and students to charter schools.

Concurrent with these reforms, while the absolute poverty rates in Newark stayed relatively constant, or increased modestly, the relative poverty rates in Newark declined as surrounding district poverty rates increased; see Figure A3 (including Orange, East Orange and Irvington). This affects any comparisons of Newark versus surrounding counterfactual districts.

Figure A3
U.S. Census Poverty Rates as a Ratio to the Mean of Districts in Newark Metropolitan Area



Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System*. Retrieved November 26, 2017, from: http://www.schoolfundingfairness.org/data-download

Appendix B
Estimating the Relative Productivity of Schools in Newark

Here, we use publicly available data to explore the following questions:

- What are the long-term trends in school performance levels on state assessments, prior to recent reforms?
- What were the shifts in school-level growth estimates that occurred during the reform period, across the test-change gap (from NJASK to PARCC), using state growth estimates and comparing against nearby comparable schools?
- Are reported "high growth" or high value-added schools identified by the authors accurately characterized as such when considering additional factors, such as available resources and economies of scale?

Performance Trends Pre-"Reform"

First, we take a brief look at pre-2010 and post-2010 changes in a) NJASK scale scores, adjusted for student population differences and b) state-calculated growth percentiles (SGP) prior to and bridging the move from the NJASK to PARCC. The reports' analysis of Newark public schools focuses only on value-added measures (notably, more thorough than growth percentiles) with two baseline years prior to the supposed reforms under investigation. Here, our intent is simply to provide some additional context regarding trends in scale scores, and potential issues arising from calculating growth across different assessments.

Comparisons against Abbott districts or all of District Factor Group (DFG) A are problematic due to regional diversity of the state in terms of demographic composition, economic conditions and neighborhood housing stock structure, quality and distribution. It's inappropriate, for example, to compare subsidized lunch rates, or any measure of "poverty" based on fixed income thresholds in Camden[9] with those of Newark or Jersey City. As Baker and colleagues show, there is substantial variation in regional income levels and costs which affect quality of life at any given income threshold.[10]

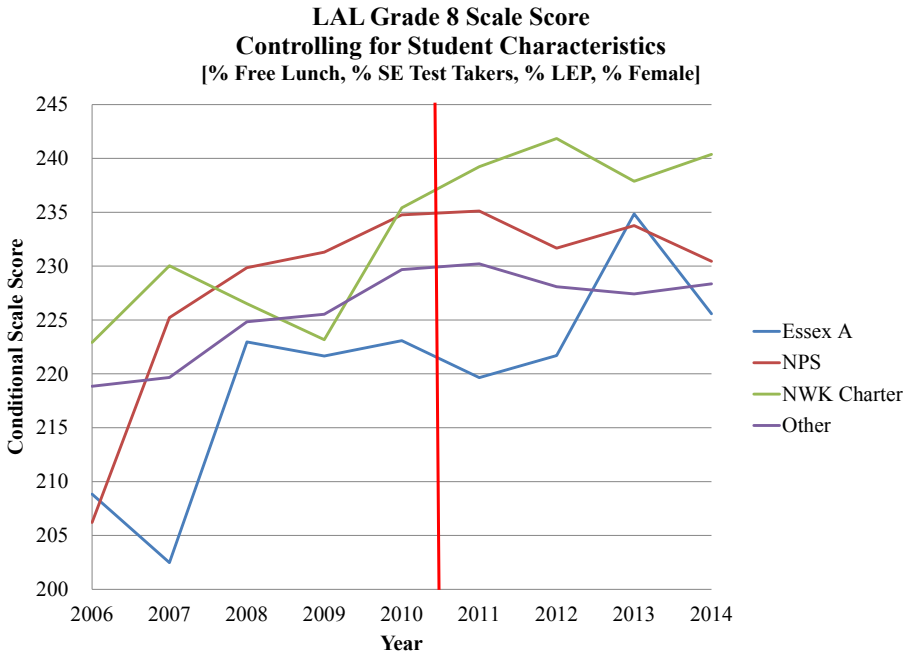
A more accurate way is to compare against schools more proximally situated. Thus, we compare Newark Public Schools with Newark Charter Schools and District Factor Group A schools in Essex County, most of which lie on the edges of Newark itself, in cities such as East Orange and Irvington.

Figure B1 and Figure B2 report regression-adjusted NJASK mean scale scores in 8th grade for a) Newark Public Schools, b) Newark Charter Schools, c) DFG A schools in other Essex County Districts, and d) other schools statewide.

Figure B1 shows that Newark district and charter schools had a generally upward trajectory of English language arts performance prior to Zuckerberg reforms, but coinciding with a

general upward drift in language arts scores statewide. Following 2010, charter and district school mean scores diverge, having converged prior to this period, which may result from policies favoring charters in more recent years, coupled with resources and practices discussed in Appendix C.

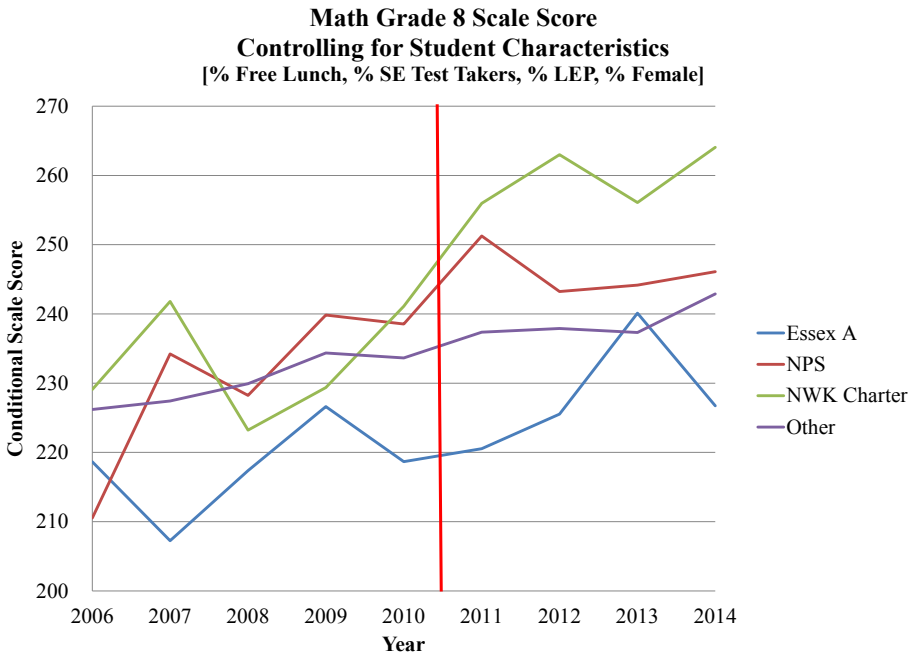
Figure B1



New Jersey Statewide Assessment Databases: http://www.nj.gov/education/schools/achievement/index.html

Figure B2 shows that Newark Charter and district math scale scores were increasing at a faster rate a) than others statewide and b) prior to 2010 more so than after 2010.

Figure B2



New Jersey Statewide Assessment Databases: http://www.nj.gov/education/schools/achievement/index.html

Growth Pre/Post-"Reform"

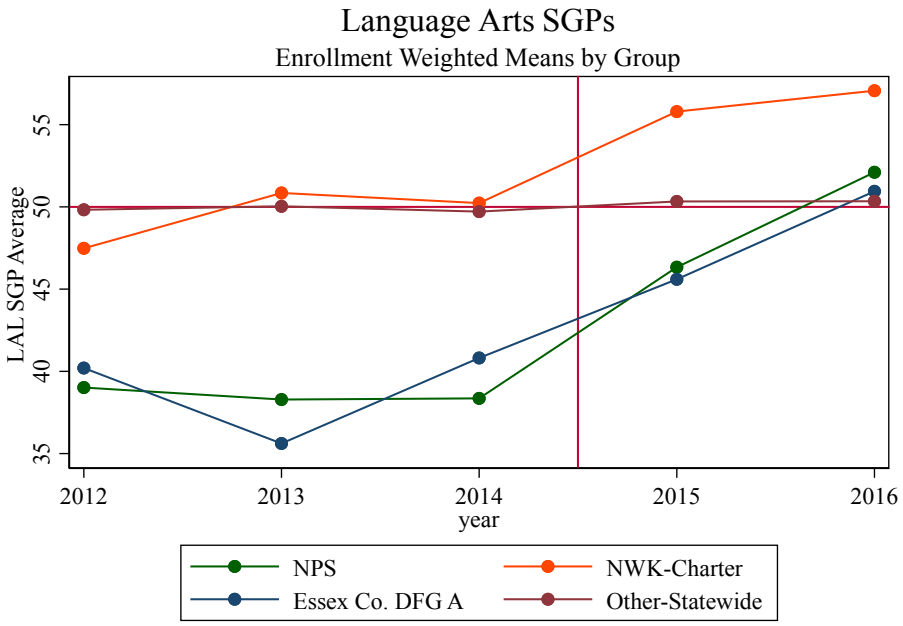
Figure B3 and Figure B4 show the enrollment weighted average of school median growth percentiles for language arts and math for a) Newark Public Schools, b) Newark Charter Schools, c) DFG A schools in other Essex County Districts and d) other schools statewide. Figure B3 establishes the relatively constant statewide (other) mean of 50 (a design feature). For the other three groups, all of which are relatively high in poverty and each of which are in Essex County, SGPs in language arts jump about 5 points between 2014 and 2015, having held more constant levels prior.

Essex DFG A schools dip then rebound between 2012 and 2014, then move parallel to Newark district and charter school between 2014 and 2015. This finding raises some suspicions regarding comparisons finding much greater growth in Newark district and charter schools when compared against more geographically and demographically diverse counterfactuals. Non-Newark DFG A schools would not have been subjected to the supposed Newark reform treatments, but seem to show similar growth between 2014 and 2015.

Figure B4 reveals similar parallels between Newark district and Essex County DFG A schools for math SGPs. Both jump between 2014 and 2015 when state tests were changed. Notably, Newark charter SGPs remained stagnant for math between these years, but at a higher level. One implication is that the test re-norming and growth calculation resulted in anomalous higher growth in previously low-growth (high-poverty) settings. Again, SGPs for other

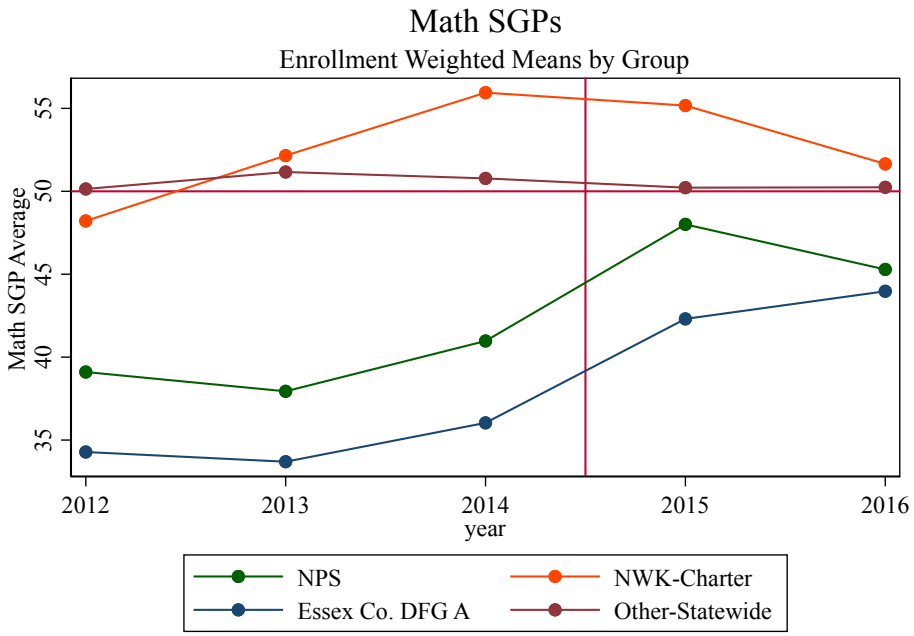
schools statewide remain (as forced by the norm-referenced calculation) around 50.

Figure B3



New Jersey School Performance Reports Databases: https://rc.doe.state.nj.us/ReportsDatabase.aspx

Figure B4



New Jersey School Performance Reports Databases: https://rc.doe.state.nj.us/ReportsDatabase.aspx

Using School Student Growth Percentiles (SGPs) to Evaluate Relative Productive Efficiency

The next several analyses explore the relative *productivity* of Newark district and charter schools. The reports purport to characterize the "productivity" of those reforms; that is, the extent to which, all else being equal, the introduction of reforms brought about by the infusion of the Zuckerberg contribution increases the relative productivity of Newark district and charter schools both collectively and separately when compared against other schools and children statewide. As noted in the previous section, it is important to consider the dynamics of those counterfactuals. Setting that question aside for the moment, we first consider what is meant by a "productivity" analysis and what methods might be used to distill changes in or differences between productivity.

The education production function considers as its dependent variable the outcomes achieved (for students) by institutions, while considering differences across institutions in controllable inputs (resources) and uncontrollable inputs (student factors, environmental factors). Given, as we have noted previously, that state and local policies have influenced the resources of Newark district schools differently from Newark's charter schools (and any other counterfactual against which we wish to compare them), it stands to reason that any legitimate productivity analysis should give consideration to resources (at least in the aggregate) and factors which may influence the costs (or production value) of those resources.

The reports, for example, assert that, on balance though not exclusively, schools that were

shuttered in Newark were schools that produced lower value-added achievement gains in math (Figure 7 in the NBER report[11], used to convey this point, shows only math gains), given student characteristics and peer characteristics, including prior achievement.[12] No consideration is given to whether schools producing higher or lower achievement gains were provided equitable resources, or operated at comparable scale.

In addition, we note that many of the measures used for characterizing students were insufficiently precise – most notably, special education classification and low-income status. We address this issue in greater detail in Appendix C. Lack of precision in the special education measure is especially problematic because it not only directly influences outcome variations but also influences budgetary pressures and resource allocation toward measured outcome goals.[13] Schools where large shares of disability populations have more severe disabilities, requiring lower case-load ratios, have less flexibility in how resources are allocated than others. Unfortunately, in our relatively simple illustrative analyses that follow, we are unable to resolve this imprecision problem and must rely on school-level aggregate disability population data (averaged over time). This imprecision creates a bias which necessarily favors Newark charter schools which serve few, if any, more severely disabled students, a topic addressed in greater detail in Appendix C.

For the following series of models, we use a panel (including only schools which exist in the panel for each year) of New Jersey school-level data for the years 2012 through 2016. These are the years for which the New Jersey Department of Education has produced school-level median growth percentile data. School-level median growth percentiles are achievement gain measures constructed using the same underlying assessment data as used by the authors in the reports. The major difference between the authors' value-added model and the state SGP measures is that the SGPs are a) not conditioned on any individual or peer characteristics other than previous test scores, and b) are not conditioned on grade levels of tested students.

We apply a series of regression models with which we assess the relative productivity – production of student achievement gains (SGPs) – for Newark district and charter schools, controlling for:

1. Student population characteristics.
2. Baseline achievement levels (school mean scale scores).
3. School grade level and size.
4. School resources (aggregate certified staffing salary expense per pupil).

Whether a charter or district school chooses to allocate their aggregate staff salary expense toward greater numbers of less experienced staff, to higher relative salaries, or toward seniority varies between charter operators and district schools (district schools being on a common salary schedule). This said, we acknowledge here our resource variable is a proxy that cannot fully account for the resource advantages some Newark schools – particularly charter schools – might enjoy relative to others.

These models are estimated on the population of Newark district and charter schools. That

is, they are regressions that describe the differences in average SGPs across schools in relation to differences in school characteristics. Because these are population data and the regressions are descriptive they merely are what they are (statistical significance in such context being inconsequential for inferential purposes, but providing some insights regarding clarity/certainty of patterns, trends or differences).

Student population characteristics include:

1. 3-year average proportion of students identified as having disabilities.
2. Percentage of students qualified for free lunch (<130% income threshold for poverty).
3. Percentage of children who are limited in their English language proficiency.

The second and third items above were obtained through statewide school-level enrollment files.[14] Special education enrollments were obtained through separate request from NJDOE. Our school-level resource measure is averaged from 2012 through 2016 (due to incompleteness of 2016 data) and includes the sum of the salaries of all certified staff (prorated for their time assigned to each site) divided by school enrollments. Staffing data, including salaries are from the statewide fall staffing report (obtained from NJDOE).

Importantly, the dependent variable is statewide and norm-referenced, where the state median is the 50th percentile and state mean roughly at 50 for SGP. As such, when we model only data on Newark schools, changes in the dependent variable are still compared against the state as a whole. We use our models to first describe variations in SGPs across schools within Newark, and then to evaluate which schools over and under perform with respect to expectations. That is, which schools really are more and less productive, given not only the students they serve, but also the resources they are provided.

Table B1 shows the regressions for elementary and middle level schools (including combined grades schools) for language arts and math student growth percentiles.

Table B1
Baseline Panel Model of SGPs with Scale Effect

	Elem Panel LAL		Elem Panel Math		MS Panel LAL		MS Panel Math	
	coef	se	coef	se	coef	se	coef	se
% Special Ed (mean over years)	-40.086*	23.785	-47.560*	28.873	-37.047*	20.467	-54.556*	28.426
% Free Lunch	-3.126	9.967	-3.220	13.746	16.062	10.388	32.711***	12.562
% ELL	-4.030	9.278	-16.982	17.989	-15.292	9.552	-33.098**	14.615
ln_salperpupil (mean over years)	20.555*	11.113	21.635	14.751	18.455*	10.191	25.398**	11.597
ln_enroll	7.064***	2.660	8.754**	4.326	5.865**	2.583	7.602**	3.266
Year (Baseline = 2012)								
2013	-0.344	1.346	-0.038	2.233	-1.828	1.736	0.920	2.981
2014	-0.975	1.562	2.252	2.400	-2.047	1.775	2.016	2.620
2015	7.683***	2.036	8.314***	2.440	7.842***	1.977	9.185***	2.278
2016	12.665***	1.517	6.088***	2.071	11.264***	1.597	6.264***	1.949
Total Mean Scale ELA	0.372***	0.128						
Total Mean Scale ELA					0.392***	0.121		
Total Mean Scale Math			0.216**	0.089				
Total Mean Scale Math							0.387***	0.061
Constant	-249.161**	124.304	-241.157	151.392	-240.655**	115.391	-323.172***	115.803
Number of observations	154		159		142		147	
R2	0.527		0.316		0.523		0.369	

note: *** p<0.01, ** p<0.05, * p<0.1

The first set of models shows that:

- Schools with more resources (higher total salaries per pupil) tend to have higher growth percentiles;
- Special education populations have substantive negative effects on growth percentiles;
- Growth dramatically jumps for schools in Newark relative to the state when shifting from NJASK to PARCC.

This third finding indicates a possibility of some anomalous affect that, we suspect, is likely an artifact of calculating growth across different tests with different distributions, under different conditions, as much if not more so than any "reforms" native to Newark exclusively.

The assertion that this large instantaneous jump is somehow a function of a treatment which suddenly kicks in at the same point in time when the testing regime is changed is at best a peculiar finding, especially since no obvious disruptive treatment occurred with timing such as to have this effect at this specific point in time. Most, if not all, reforms and non-reform drifts/changes in programs, practices and resources a) occur more gradually over time, and b) likely have longer lagged effects and less "shock-like" effects on student outcome measures.

We can use the residuals of the above models to identify which schools systematically exceed growth expectations and which schools systematically underperform. A core assertion of the reports is that the positive effects of reforms in Newark result largely from moving students out of low value-added schools and into high value-added schools, closing low value-added schools in the process. The reports further assert that Newark's charter sector is particularly effective in its production of value-added, especially the larger established operators such as

North Star and TEAM Academy (a KIPP affiliate).

Table B2 shows the residuals of our model estimated using all years, separate models for individual years, and the mean of the year-by-year models. Collapsing Newark schools to their district average, we see that Newark district schools fall around the middle of the pack, with marginally below expected growth year after year (relative within year position, among schools in Newark only). The residual values in the table indicate, for example, that MV Rogers Charter School's actual SGPs were 7.13 higher than expected, on a scale of 0 to 100, with a statewide mean (and median, forcibly normally distributed) of approximately 50. North Star Academy falls just below Newark Public schools, but has volatile shifts in growth relative to expectations year over year. TEAM falls at the bottom of the list, with consistently "below expected" performance.

Table B2
Model residuals and ranking including scale term (Charters vs. District)

School/ District	Panel Model	Mean of Year by Year	2016	2015	2014	2013	2012
Maria Varisco Rogers	7.13	6.76	6.71	13.56	3.88	9.40	1.86
New Horizons Community	5.57	4.67	11.20	1.98	9.46	12.19	(11.47)
Marion P. Thomas	2.22	2.17	6.35	(7.01)	5.85	3.01	1.28
Gray	2.11	1.60	5.14	5.46	(4.84)	(0.01)	6.89
Newark Educators Community	1.14	1.17	5.50	(18.65)	9.10	18.29	(8.39)
NEWARK PUBLIC SCHOOLS	(1.82)	(2.37)	(0.63)	(1.84)	(1.87)	(2.73)	(2.09)
North Star Academy	(3.65)	(5.79)	(17.90)	6.52	0.89	(12.96)	(2.53)
Robert Treat Academy	(4.19)	(3.97)	(7.37)	0.78	1.81	(3.18)	(12.87)
University Heights	(5.00)	(4.12)	(5.89)	0.13	(9.23)	(21.09)	14.17
TEAM Academy	(9.13)	(8.10)	(12.57)	(10.68)	(8.19)	(5.65)	(4.56)

Notably there is significant variation among Newark schools, including some very high and some very low performers, and others for which there exists significant year over year volatility. Table B3 lists individual school rankings. The very small Discovery Charter schools ranks highest, but possibly because the scale measure in our model plays to their advantage. Other Newark district K-8 schools rank quite high, and as noted in Table B2, TEAM Academy ranks quite low. North Star is nearer the middle when adjusting for the factors in the model.

Table B3
School Rankings including Scale

School	Grade Span	Panel Rating	Mean Across Years	2016	2015	2014	2013	2012
DISCOVERY CS	04-08	22.59	28.36	19.36	-0.76	12.69	33.22	51.13
MT VERNON PLACE SCHOOL	PK-08	9.04	9.33	10.67	9.34	15.22	5.67	6.05
FIRST AVENUE SCHOOL	PK-08	8.61	8.81	2.79	12.97	7.92	10.28	8.70
CHANCELLOR AVENUE SCHOOL	03-08	7.51	8.58	5.09	4.86	2.73	5.89	20.80
MARIA L. VARISCO-ROGERS	KG-08	7.13	6.76	6.71	13.56	3.88	9.40	1.86
HAWTHORNE AVENUE SCHOOL	KG-08	5.93	8.35	5.37	-8.88	0.11	18.77	14.55
NEW HORIZONS COMM. CS	KG-05	5.57	4.67	11.20	1.98	9.46	12.19	-11.47
ELLIOTT STREET ELEMENTARY SCHOOL	PK-04	4.51	5.10	-13.23	-11.87	24.66	11.99	13.95
CAMDEN STREET ELEMENTARY SCHOOL	PK-08	4.27	5.44	2.92	3.15	7.02	7.23	4.18
IVY HILL ELEMENTARY SCHOOL	PK-08	4.07	5.12	5.36	0.17	7.94	3.83	5.72
SUSSEX AVENUE SCHOOL Burnet Street School	PK-08	3.96	2.98	8.09	5.12	2.62	-6.32	9.03
HAWKINS STREET SCHOOL	PK-08	3.43	3.31	-1.37	9.27	10.69	3.13	-5.17
WILSON AVENUE SCHOOL	PK-08	2.40	3.80	2.84	-2.54	4.45	8.09	1.00
MARION P. THOMAS CS	PK-08	2.22	2.17	6.35	-7.01	5.85	3.01	1.28
GRAY CS	KG-08	2.11	1.60	5.14	5.46	-4.84	-0.01	6.89
Dr. E. ALMA FLAGG SCHOOL	KG-08	1.94	2.60	6.90	-3.95	0.90	-0.43	8.53
DAYTON STREET SCHOOL at Peshine Avenue	PK-08	1.90	2.14	5.33	-2.00	3.61	-0.62	3.74
MCKINLEY	PK-08	1.23	0.93	4.89	1.30	-1.14	-1.13	2.24
NEWARK EDUCATORS CHARTER	KG-05	1.14	1.17	5.50	-18.65	9.10	18.29	-8.39
BENJAMIN FRANKLIN ELEMENTARY SCHOOL	PK-04	0.57	2.15	5.87	-1.13	-0.73	11.24	-4.52
ROBERTO CLEMENTE ELEMENTARY SCHOOL	PK-04	0.40	1.23	6.32	3.18	-0.81	2.43	-4.99
MILLER STREET SCHOOL	PK-08	0.37	2.46	-8.94	1.96	1.83	-1.72	12.71
QUITMAN COMMUNITY SCHOOL	PK-08	0.33	-4.02	1.90	6.86	0.52	-17.36	-4.36
SOUTH SEVENTEENTH STREET SCHOOL	KG-08	0.08	0.39	-3.60	-2.99	0.43	-2.39	7.56
LOUISE A SPENCER ELEMENTARY SCHOOL	PK-08	-0.48	-2.18	0.78	-3.85	-16.13	3.61	0.73
BELMONT RUNYON ELEMENTARY SCHOOL	PK-07	-0.76	-1.92	-1.38	-3.22	-1.81	-6.14	0.73
LAFAYETTE STREET SCHOOL	PK-08	-0.87	-1.05	1.97	-2.25	-2.96	-1.66	-0.32
ABINGTON AVENUE SCHOOL	PK-08	-1.28	-1.87	7.56	-1.07	0.96	1.90	-14.28

School	Grade Span	Panel Rating	Mean Across Years	2016	2015	2014	2013	2012
THIRTEENTH AVENUE SCHOOL MARTIN LUTHER KING	PK-08	-1.67	-2.90	-1.25	-0.77	-4.04	-5.84	-1.16
ANN STREET SCHOOL	PK-08	-1.92	-1.85	0.17	1.75	-5.65	-7.75	4.85
SPEEDWAY AVENUE SCHOOL	PK-08	-2.07	-2.34	10.92	-3.11	0.05	-3.75	-8.98
BRANCH BROOK SCHOOL	PK-04	-2.53	-2.34	-5.20	-22.32	0.04	1.23	14.55
DR WILLIAM H HORTON ELEMENTARY SCHOOL	KG-08	-3.27	-1.85	-9.46	-4.92	-0.40	6.36	-5.64
HARRIET TUBMAN ELEMENTARY SCHOOL	PK-06	-3.29	-2.98	-4.70	-1.07	-12.18	3.07	0.00
NORTH STAR ACAD. CS OF N	KG-12	-3.65	-5.79	-17.90	6.52	0.89	-12.96	-2.53
LINCOLN	PK-08	-3.99	-3.68	-6.33	-7.19	-5.97	0.39	-0.29
ROBERT TREAT ACADEMY CS	KG-08	-4.19	-3.97	-7.37	0.78	1.81	-3.18	-12.87
OLIVER STREET SCHOOL	PK-08	-4.21	-2.89	0.40	-13.65	-6.65	1.20	1.94
UNIVERSITY HEIGHTS CS	KG-05	-5.00	-4.12	-5.89	0.13	-9.23	-21.09	14.17
GEORGE WASHINGTON CARVER ELEMENTARY SCHOOL	PK-08	-5.07	-5.51	-4.93	-1.14	-5.00	-11.52	-1.43
RAFAEL HERNANDEZ SCHOOL	PK-08	-5.60	-6.38	-7.22	1.26	5.08	-7.93	-19.11
American History High School	07-12	-5.94	-19.63	5.03	-10.87	-5.86	-4.17	-48.86
CLEVELAND Eighteenth Avenue School	PK-08	-6.56	-9.73	5.94	-3.40	-14.84	-16.87	-8.44
ARTS HIGH SCHOOL	07-12	-6.92	-9.81	-10.51	5.33	-12.42	-14.32	-2.71
AVON AVENUE SCHOOL	KG-08	-7.12	-6.99	-13.21	-6.28	-11.46	-4.94	-0.45
SOUTH STREET ELEMENTARY SCHOOL	PK-05	-7.84	-6.90	-5.79	-16.04	-6.53	-3.35	-2.80
TEAM ACADEMY CHARTER SCH	KG-12	-9.13	-8.10	-12.57	-10.68	-8.19	-5.65	-4.56
RIDGE STREET SCHOOL	KG-08	-11.82	-12.22	-3.04	-16.26	-19.56	-7.35	-11.39
UNIVERSITY HIGH SCHOOL	07-12	-20.80	-23.53	-16.51	-7.74	-16.02	-26.18	-28.38
SCIENCE PARK HIGH SCHOOL	07-12	-27.76	-35.70	-18.52	-16.26	-22.54	-47.34	-37.22

Unconditional growth percentiles for TEAM are relatively average and raw growth percentiles for North Star tend to be quite high, and consistently so, causing us to dig deeper into these findings. The model in Table B1 finds that on average, larger schools tend to have higher SGPs. North Star and TEAM are essentially districts operating within the Newark district, but are reported as schools in state data. By comparison to other schools, they are very large. Using the models in Table B1, where larger schools are expected to have higher SGPs ends up setting a very high "expected" bar for North Star and TEAM.

As such, we estimate a second set of models and residuals removing the school size measure. Table B4 shows the model including district and charter schools with the school size measure removed. The removal of the size measure compromises the statistical significance of

the relationship between resources and outcomes, cutting the magnitude of this relationship in half (because of the way in which staffing ratios and total salary expense interact with school size). We still, however, have the anomalous jump in SGPs in 2015, the year of the switch from the NJASK to the PARCC. Our model here also explains less overall variation in SGPs (as compared to models including enrollment size).

Table B4
Panel Model with Scale Effect Excluded

	Elem Panel LAL		Elem Panel Math		MS Panel LAL		MS Panel Math	
	coef	**se**	**coef**	**se**	**coef**	**se**	**coef**	**se**
% Special Ed (mean over years)	-19.197	19.013	-20.838	23.790	-19.105	16.778	-31.494	22.648
% Free Lunch	-7.076	10.012	-6.385	14.044	9.881	9.633	29.295**	12.166
% ELL	8.680	8.527	-1.396	16.190	-3.542	8.706	-19.401	13.344
ln_salperpupil (mean over years)	8.670	10.914	7.698	14.244	8.577	10.021	14.185	11.280
Year (Baseline = 2012)								
2013	0.190	1.365	0.668	2.172	-1.206	1.750	1.583	2.994
2014	-0.835	1.564	2.573	2.391	-1.952	1.740	2.350	2.583
2015	**7.493*** **	**1.990**	**8.288*** **	**2.305**	**7.563*** **	**1.924**	**9.215*** **	**2.199**
2016	**13.533*** **	**1.585**	**7.334*** **	**1.993**	**12.035*** **	**1.670**	**7.491*** **	**1.864**
Total Mean Scale ELA	0.347***	0.131						
Total Mean Scale ELA					0.366***	0.126		
Total Mean Scale Math			0.217**	0.087				
Total Mean Scale Math							0.395***	0.066
Constant	-94.447	116.958	-64.008	137.463	-109.061	109.012	-177.990*	102.271
Number of observations	154		159		142		147	
R2	0.496		0.276		0.499		0.333	

note: *** p<0.01, ** p<0.05, * p<0.1

We again calculate the residuals from this model without scale and present them in Table B5. Significant changes are apparent compared with residuals from the earlier model (Table B2). Here, TEAM academy is relatively average, alongside (marginally above in some years) Newark's district schools.

North Star jumps to the top of the list, but notably in some years (2013 & 2016) is only slightly above the middle. Given these models, one might consider North Star a high value-added producer. But, as we discuss in Appendix C, these models fail to capture some unique features of the "North Star Model."

Table B5
Model residuals and ranking EXCLUDING scale term (Charters vs. District)

School/ District	Panel Model	Mean of Year by Year	2016	2015	2014	2013	2012
North Star Academy	7.99	6.79	1.92	10.85	7.66	0.61	15.34
Maria Varisco Rogers	6.55	6.12	3.62	13.38	3.61	8.29	2.29
Discovery	6.01	11.29	(15.80)	(7.61)	2.88	13.62	33.52
New Horizons Community	3.61	2.93	5.53	0.61	7.84	10.35	(9.70)
Marion P. Thomas	2.94	3.15	5.44	(6.82)	6.78	3.46	4.14
TEAM Academy	0.23	2.23	3.32	(7.14)	(2.41)	5.00	9.68
THE NEWARK PUBLIC SCHOOLS	(1.70)	(2.50)	(1.39)	(1.69)	(1.95)	(2.88)	(2.02)
University Heights	(3.85)	(3.06)	(5.38)	0.51	(8.80)	(19.94)	17.46
Newark Educators Community	(4.07)	(4.47)	(5.40)	(20.85)	5.55	12.76	(14.40)
Robert Treat Academy	(5.87)	(5.68)	(13.29)	0.13	1.36	(5.67)	(13.69)
The Gray	(6.53)	(7.32)	(14.65)	1.95	(9.21)	(10.81)	(1.95)

Table B6 shows that while North Star ranks high among individual schools, two Newark district schools still rank higher.

Table B6
School Rankings Excluding Scale

School	Grade Span	Panel Rating	Mean Across Years	2016	2015	2014	2013	2012
MT VERNON PLACE SCHOOL	PK-08	11.49	11.17	10.89	13.41	16.52	8.40	7.92
FIRST AVENUE SCHOOL	PK-08	9.83	10.01	4.58	13.19	8.65	10.91	10.81
NORTH STAR ACAD. CS OF N	KG-12	9.07	7.29	0.43	14.58	8.22	2.19	14.62
MARIA L. VARISCO-ROGERS	KG-08	7.66	6.96	2.89	16.05	4.67	9.73	2.85
CHANCELLOR AVENUE SCHOOL	03-08	5.48	6.53	-1.08	3.14	1.59	3.97	18.79
DISCOVERY CS	04-08	5.15	9.80	-18.28	-8.02	4.13	12.22	28.77
HAWTHORNE AVENUE SCHOOL	KG-08	4.28	6.76	-1.14	-9.02	-0.11	16.86	13.67
DAYTON STREET SCHOOL at Peshine Avenue	PK-08	4.10	5.23	6.81	0.09	5.99	2.91	7.55
QUITMAN COMMUNITY SCHOOL	PK-08	3.80	-1.22	4.24	9.08	2.47	-14.04	-0.21
WILSON AVENUE SCHOOL	PK-08	3.78	4.90	3.31	-0.96	5.22	9.54	2.18
MCKINLEY	PK-08	3.42	3.38	7.93	1.13	0.32	1.60	5.66
ELLIOTT STREET ELEMENTARY SCHOOL	PK-04	3.28	3.53	-18.08	-10.94	23.09	11.77	11.81
IVY HILL ELEMENTARY SCHOOL	PK-08	2.97	3.60	2.24	0.57	7.15	1.57	4.18
NEW HORIZONS COMM. CS	KG-05	2.17	1.58	4.75	0.83	6.55	8.56	-12.80
CAMDEN STREET ELEMENTARY SCHOOL	PK-08	1.90	3.01	-1.50	-0.25	5.40	4.71	1.08

School	Grade Span	Panel Rating	Mean Across Years	2016	2015	2014	2013	2012
HAWKINS STREET SCHOOL	PK-08	1.54	1.00	-5.58	8.41	9.31	0.47	-8.21
SUSSEX AVENUE SCHOOL Burnet Street School	PK-08	1.13	-0.10	2.14	4.33	0.94	-10.58	6.05
MARION P. THOMAS CS	PK-08	0.90	1.03	5.02	-9.13	5.05	0.12	1.44
American History High School	07-12	0.88	-14.57	7.72	-2.34	-1.41	1.43	-43.73
BELMONT RUNYON ELEMENTARY SCHOOL	PK-07	0.31	-0.24	-1.00	-1.97	-0.09	-4.78	3.92
LOUISE A SPENCER ELEMENTARY SCHOOL	PK-08	-0.36	-1.44	-0.09	-3.58	-14.88	5.17	1.78
TEAM ACADEMY CHARTER SCH	KG-12	-0.52	1.26	2.53	-7.43	-3.12	3.81	7.84
Dr. E. ALMA FLAGG SCHOOL	KG-08	-0.57	0.04	1.74	-4.98	-0.50	-3.71	5.82
SOUTH SEVENTEENTH STREET SCHOOL	KG-08	-0.67	0.12	-7.16	-4.03	0.44	-2.72	8.47
THIRTEENTH AVENUE SCHOOL MARTIN LUTHER KING	PK-08	-0.80	-1.38	1.27	-2.09	-3.11	-4.79	1.37
LAFAYETTE STREET SCHOOL	PK-08	-1.39	-1.53	0.82	-3.49	-3.12	-2.87	-0.29
UNIVERSITY HEIGHTS CS	KG-05	-1.56	-1.61	-6.45	3.71	-6.82	-17.46	18.29
ARTS HIGH SCHOOL	07-12	-1.97	-6.74	-3.61	12.48	-10.41	-10.31	0.50
MILLER STREET SCHOOL	PK-08	-2.39	-0.48	-15.70	0.84	0.08	-4.52	8.97
ABINGTON AVENUE SCHOOL	PK-08	-2.49	-3.15	4.60	-1.94	0.31	0.30	-15.73
ANN STREET SCHOOL	PK-08	-2.86	-3.30	-0.71	0.97	-6.86	-9.91	2.75
DR WILLIAM H HORTON ELEMENTARY SCHOOL	KG-08	-3.05	-1.60	-8.89	-4.65	-0.43	6.26	-5.13
HARRIET TUBMAN ELEMENTARY SCHOOL	PK-06	-3.08	-3.15	-9.87	0.82	-12.17	4.34	1.11
GEORGE WASHINGTON CARVER ELEMENTARY SCHOOL	PK-08	-3.46	-4.21	-4.04	0.22	-4.20	-10.71	1.12
ROBERTO CLEMENTE ELEMENTARY SCHOOL	PK-04	-3.57	-3.18	-0.14	1.71	-4.00	-1.54	-11.94
NEWARK EDUCATORS CHARTER	KG-05	-4.14	-4.50	-6.17	-19.99	5.15	13.16	-14.65
BENJAMIN FRANKLIN ELEMENTARY SCHOOL	PK-04	-4.68	-3.92	-2.24	-3.39	-4.71	5.94	-15.18
BRANCH BROOK SCHOOL	PK-04	-4.86	-5.10	-17.11	-19.76	-3.26	1.68	12.93
LINCOLN	PK-08	-5.38	-5.24	-10.63	-6.80	-6.91	-1.93	-1.24
SPEEDWAY AVENUE SCHOOL	PK-08	-5.59	-5.52	5.50	-4.98	-1.45	-8.91	-11.89
ROBERT TREAT ACADEMY CS	KG-08	-5.67	-6.35	-13.97	1.96	0.99	-6.29	-15.34
OLIVER STREET SCHOOL	PK-08	-6.17	-5.22	-3.20	-14.13	-8.64	-1.21	-1.42
AVON AVENUE SCHOOL	KG-08	-6.45	-5.90	-13.20	-6.33	-10.76	-4.65	2.75
CLEVELAND Eighteenth Avenue School	PK-08	-7.09	-10.68	1.30	-3.49	-15.09	-18.00	-8.76
RAFAEL HERNANDEZ SCHOOL	PK-08	-7.52	-8.21	-10.51	-1.16	3.91	-10.66	-20.51
GRAY CS	KG-08	-8.68	-10.21	-15.53	1.08	-11.46	-14.84	-6.44
RIDGE STREET SCHOOL	KG-08	-10.69	-11.66	-5.14	-13.00	-19.20	-6.08	-10.84
SOUTH STREET ELEMENTARY SCHOOL	PK-05	-12.08	-11.75	-15.19	-16.77	-10.04	-6.95	-9.80

School	Grade Span	Panel Rating	Mean Across Years	2016	2015	2014	2013	2012
UNIVERSITY HIGH SCHOOL	07-12	-15.40	-19.05	-10.00	-1.67	-12.82	-20.70	-23.61
SCIENCE PARK HIGH SCHOOL	07-12	-15.91	-25.36	-8.01	-1.88	-14.22	-33.60	-28.25

To summarize the findings of the two alternative specifications:

- TEAM is either the lowest producer of growth, a relatively average producer of growth over time, or somewhere in between.
- North Star is either a relatively average producer of growth, one of the highest producers of growth over time, or somewhere in between.
- Several NPS district schools are high performers in terms of production of annual student growth.
- There is significant volatility in these models from year to year.

We also tested to see if the size effect found in our first set of models was driven primarily by the large size and high performance of North Star and TEAM; in other words, if we estimated the same model to only Newark District schools, would we still find a large positive size effect in relation to SGPs? Table B7 shows those regressions and reveals that the size effect is robust to the exclusion of charter schools. Further, including the size effect in this case, again, adds clarity to the resource-outcome relationships. That is, there's something to the relationship between school size and outcomes in Newark (though most likely in a more complex and non-linear relationship than our log-linear transformation). Tests of alternative functional forms did not reveal additional insights.

Table B7
Robustness Check: Is the positive influence of scale driven by TEAM and North Star? (Model Excluding Charters)

	Elem Panel LAL		Elem Panel Math		MS Panel LAL		MS Panel Math	
	coef	se	coef	se	coef	se	coef	se
% Special Ed (mean over years)	-34.124	22.669	-47.760	30.466	-32.561*	18.570	-60.140**	30.355
% Free Lunch	-5.488	13.465	-8.757	18.383	12.007	13.499	31.056**	14.897
% ELL	-4.896	11.629	-19.353	21.956	-17.242	11.334	-39.083**	18.818
ln_salperpupil (mean over years)	19.388	12.136	21.910	15.810	19.972*	11.009	31.493**	13.100
ln_enroll	**8.867*** **	**3.130**	**11.044** **	**5.282**	**7.185** **	**3.232**	**7.817***	**4.269**
Year (Baseline = 2012)								
2013	-1.269	1.442	-1.195	2.553	-3.003*	1.750	-0.010	3.360
2014	-1.980	1.712	0.817	2.611	-3.326*	1.818	0.858	2.804
2015	6.891***	2.020	8.670***	2.562	6.914***	1.950	9.907***	2.321
2016	12.394***	1.703	6.014***	2.240	10.810***	1.765	6.633***	2.108
Total Mean Scale ELA	0.324**	0.135						
Total Mean Scale ELA					0.406***	0.147		
Total Mean Scale Math			0.183	0.114				
Total Mean Scale Math							0.455***	0.099
Constant	-240.036*	133.369	-246.756	159.462	-261.819**	125.180	-388.422***	130.163
Number of observations	135		140		121		126	
R2	0.568		0.328		0.587		0.408	

note: *** p<0.01, ** p<0.05, * p<0.1

Differential Growth in Charter and District Schools Over Time

Finally, we estimated a series of models to try to tease out the relative shifts in growth percentiles over time in charter and district schools in Newark. Table B8 shows the models of SGPs for Newark charter and district schools with a) the size term included, and with b) interactions between charter status and year. To summarize our findings:

- Again, resources are positively associated with SGPs.
- Again, we have what appears to be an anomalous, large jump in SGPs in 2015 and 2016 over their 2012 levels (as baseline year).
- Charter schools on average and across years (main effect) do not differ from district schools in their SGPs.
- Relative to baseline year, for charters, we have higher growth in 2013 and 2014, but not in 2015 and 2016, when SGPs jumped across the board.

Again, the jump in SGPs observed is large. Because these are state normed measures, this finding does indicate a large jump in SGPs relative to statewide SGP (which remain stable around 50). But again, this jump is:

- sudden,
- large,
- timed with change in test, and
- not timed with any clear, obvious, large disruptive innovation.

It is difficult at best to swallow the premise that ill-defined, subtle policy changes and ongoing enrollment shifts which occurred during this period had a sudden and large effect on achievement growth, which just so happened to coincide with a change in test.

Table B8
Model including charter effect and charter by year effect (with scale effect)

	Elem Panel LAL		Elem Panel Math		MS Panel LAL		MS Panel Math	
	coef	**se**	**coef**	**se**	**coef**	**se**	**coef**	**se**
% Special Ed (mean over years)	**-42.034***	23.202	**-50.539***	28.563	**-39.229****	19.663	**-56.332****	28.713
% Free Lunch	0.605	12.201	-1.032	14.678	19.563*	11.434	32.623**	12.793
% ELL	1.538	13.133	-16.866	22.465	-10.016	12.398	-33.878*	18.796
ln_salperpupil (mean over years)	**23.462****	11.469	**23.271***	13.508	**22.176****	10.613	**26.126****	10.783
ln_enroll	8.160***	2.675	10.232***	3.819	7.075***	2.455	8.626***	2.907
Year (Baseline = 2012)								
2013	-1.344	1.463	-1.216	2.447	-3.225*	1.753	-0.238	3.320
2014	-1.927	1.724	0.976	2.558	-3.394*	1.810	0.491	2.776
2015	**6.996*****	2.001	**9.010*****	2.540	**6.926*****	1.904	**9.508*****	2.331
2016	**12.540*****	1.728	**6.323*****	2.208	**10.806*****	1.830	**6.215*****	2.093
Charter	0.979	5.584	-0.001	5.789	-2.106	5.967	-2.879	6.274
(year==2013)*charter	7.348	5.457	**10.897*****	3.073	**11.377****	5.708	**9.698****	4.523
(year==2014)*charter	**7.452****	3.714	10.550	7.201	**10.845****	4.286	**12.600***	7.595
(year==2015)*charter	4.365	8.488	-5.585	9.705	6.741	8.083	-2.343	9.670
(year==2016)*charter	-0.786	2.615	-3.782	3.896	2.711	3.233	-0.801	3.915
Total Mean Scale ELA	0.330**	0.129						
Total Mean Scale ELA					0.372***	0.114		
Total Mean Scale Math			0.202**	0.103				
Total Mean Scale Math							0.376***	0.076
Constant	-277.371**	127.445	-263.715*	137.472	-280.276**	118.208	-333.355***	104.312
Number of observations	154		159		142		147	
R2	0.548		0.354		0.551		0.402	

note: *** p<0.01, ** p<0.05, * p<0.1

Table B9 takes another run at these data, excluding the size term, leading to largely the same findings (but for mitigation of the relationship between resources and outcomes). The main effect for charters does not change substantively, nor do the charter-by-year interactions. Charters show some positive differences in SGPs in 2013 and 2014, but not in 2015 and 2016.

Table B9
Model including charter effect and charter by year effect (without scale effect)

	Elem Panel LAL		Elem Panel Math		MS Panel LAL		MS Panel Math	
	coef	**se**	**coef**	**se**	**coef**	**se**	**coef**	**se**
% Special Ed (mean over years)	-18.578	18.275	-20.254	24.928	-18.570	16.062	-31.556	24.341
% Free Lunch	-6.496	11.977	-7.698	15.504	10.188	10.782	26.677**	12.718
% ELL	12.746	12.902	-3.480	21.963	-0.235	11.787	-23.813	18.400
ln_salperpupil (mean over years)	8.982	11.839	6.308	14.410	9.343	11.099	12.565	11.618
Year (Baseline = 2012)								
2013	-0.807	1.472	-0.401	2.378	-2.531	1.735	0.539	3.326
2014	-1.843	1.713	1.374	2.539	-3.274*	1.747	0.942	2.727
2015	**6.432*** **	**1.966**	**8.679*** **	**2.494**	**6.376*** **	**1.876**	**9.374*** **	**2.325**
2016	**13.190*** **	**1.810**	**7.458*** **	**2.232**	**11.433*** **	**1.896**	**7.316*** **	**2.070**
Charter	-3.055	6.665	-4.800	6.932	-5.875	7.037	-7.224	7.496
(year==2013)*charter	8.796	5.357	12.109***	2.863	12.427**	5.525	10.686**	4.427
(year==2014)*charter	8.394**	3.655	11.214	7.271	11.181***	4.042	12.568	7.784
(year==2015)*charter	7.575	7.439	-2.030	8.352	9.124	7.267	0.176	8.642
(year==2016)*charter	2.509	3.204	-0.123	3.598	5.520	3.730	2.347	4.518
Total Mean Scale ELA	0.311**	0.133						
Total Mean Scale ELA					0.353***	0.118		
Total Mean Scale Math			0.215**	0.101				
Total Mean Scale Math							0.398***	0.085
_cons	-90.915	123.954	-49.621	140.035	-113.322	117.171	-161.018	106.458
Number of observations	154		159		142		147	
R2	0.510		0.305		0.520		0.361	

note: *** p<0.01, ** p<0.05, * p<0.1

Appendix C
Beneath the Veil of Newark's Charter Productivity Claims

Among the take-away points of the previous sections are that:

- Resources, when considering school size, are positively associated with growth;
- The productivity of large charter operators in Newark – TEAM and North Star in particular – depends on how we treat school size in our models;
- Jumps in student growth percentiles across the board between 2014 and 2015 are hard to explain as a function of substantive policy change. Policy and contextual changes had been happening gradually prior to and throughout the period.

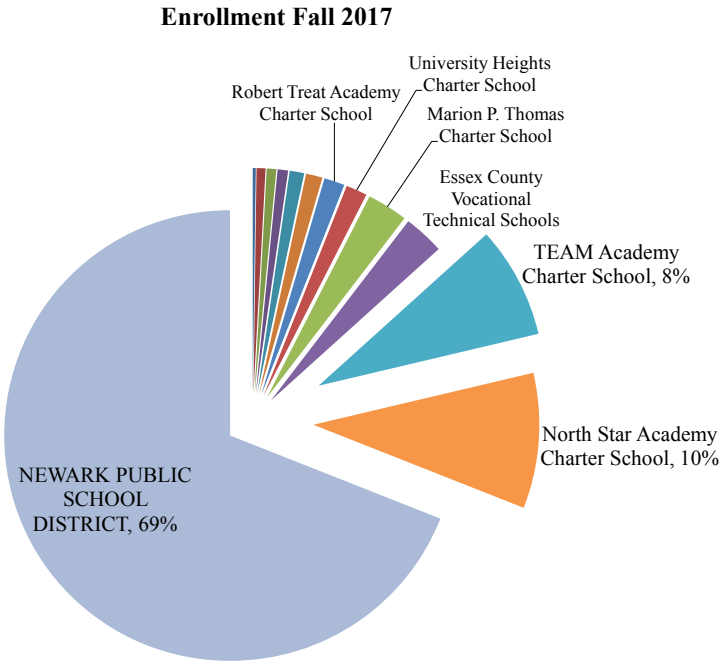
From any study of the effects of changes in policy and practices on student outcomes, what we really want to know – where positive outcome effects are observed – is what can be done to distribute those positive effects across more children and settings, and/or yield even stronger positive effects.

The conclusion offered in the reports is that shifting students to higher value-added schools has yielded positive growth in language arts. And thus, the logical policy conclusion is that more students should be shifted to high value-added schools. The larger the share of students placed in these schools, the higher the overall system performance will be. This may be an oversimplification, but is certainly the message that some are taking home from the reports.[15]

Figure C1 shows the present distribution of students across district and charter schools within the city of Newark. One might characterize the system as housing three separate K-12 school districts with a handful of smaller operators of select grade-level schools. The three comprehensive districts in question are NPS, TEAM and North Star. Analyses in the previous section (setting aside the scale question) suggest that TEAM and NPS perform similarly and that North Star tends to be the higher producer of student growth. Thus, the assertion would be that if we shift more students into North Star, more students should be better off and the system as a whole should produce better outcomes on average.

Thus the "between-school" treatment here is essentially defined as "North Starring" more students. But what exactly does that mean? Here, we attempt to provide some relevant context. Our intent is to separate the treatment of "North Starring" into those actions district leaders and policymakers might take which are desirable and scalable, versus those practices and conditions that are likely to be influencing measured outcomes but may not be scalable or desirable.

Figure C1
Distribution of District and Charter School Enrollments in Newark 2017



Source: New Jersey Department of Education, Enrollment files, 2016-17.
http://www.state.nj.us/education/data/enr/enr17/

Student Population Differences

Unfortunately, a consistent feature of North Star Academy over time has been the tendency to serve and retain less needy student populations than the broader population in the district as well as other charter operators including TEAM. Neither TEAM nor North Star serve many children with severe disabilities, but North Star serves very few with disabilities of any degree of severity. The reports' analysis fails to parse severity of disability – its influence on individual student growth, the potential peer effects of the presence of children with severe disabilities, or the extent to which larger shares of children with severe disabilities create resource allocation constraints and pressures in schools. This is a substantial omission, but one which could not be remedied given the lack of data precision.

North Star has also consistently served fewer of the lowest income children. Again, the reports' analysis fails to parse income levels across children, using only indicators of children qualified for either free or reduced priced lunch. We provide illustrations in this section demonstrating why this matters.

North Star serves effectively no children with limited English language proficiency, in part because North Star caters to a predominantly black student population from Newark's black

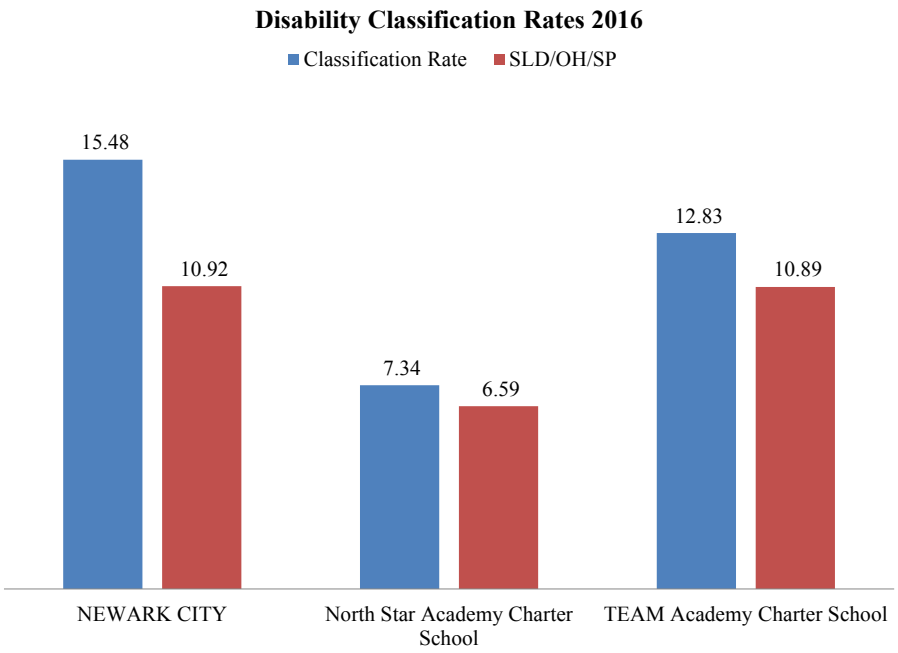
neighborhoods, which remain geographically segregated from the city's Hispanic and other ethnic neighborhoods and are home to non-English speaking families.

Special Education Rates

We start with disability rates based on 2016 data, which are actually more similar across the three Newark districts than prior years during the period studied. Figure C2 shows the overall percent classified and percent with mild specific learning disability, other health impairment, or speech/language disability. Newark Public Schools has an overall rate higher than either of the other two and more than double that of North Star. The vast majority of children with disabilities in North Star have relatively mild and less-costly disabilities. The case is similar for TEAM. Notably, TEAM and NPS have similar rates of mild disability students, but NPS has far more severe disability students.

This finding actually serves to rebut a common argument of charter advocates regarding their lower disability classification rates. Charter advocates frequently assert that effective early grades interventions reduce their need to classify students with disabilities.[16] But even the most effective interventions would only be successful at reducing the number of children identified as having mild specific learning disabilities – children on the margins of classification. Interventions would be far less likely to reduce classification of children with traumatic brain injury, intellectual disability, emotional disturbance, or autism. It is those more severe and costly disabilities which are more prevalent in the NPS schools. Whether valid in other settings or not, this argument is unlikely to hold for differences in special education classification rates between NPS and TEAM Academy.

Figure C2

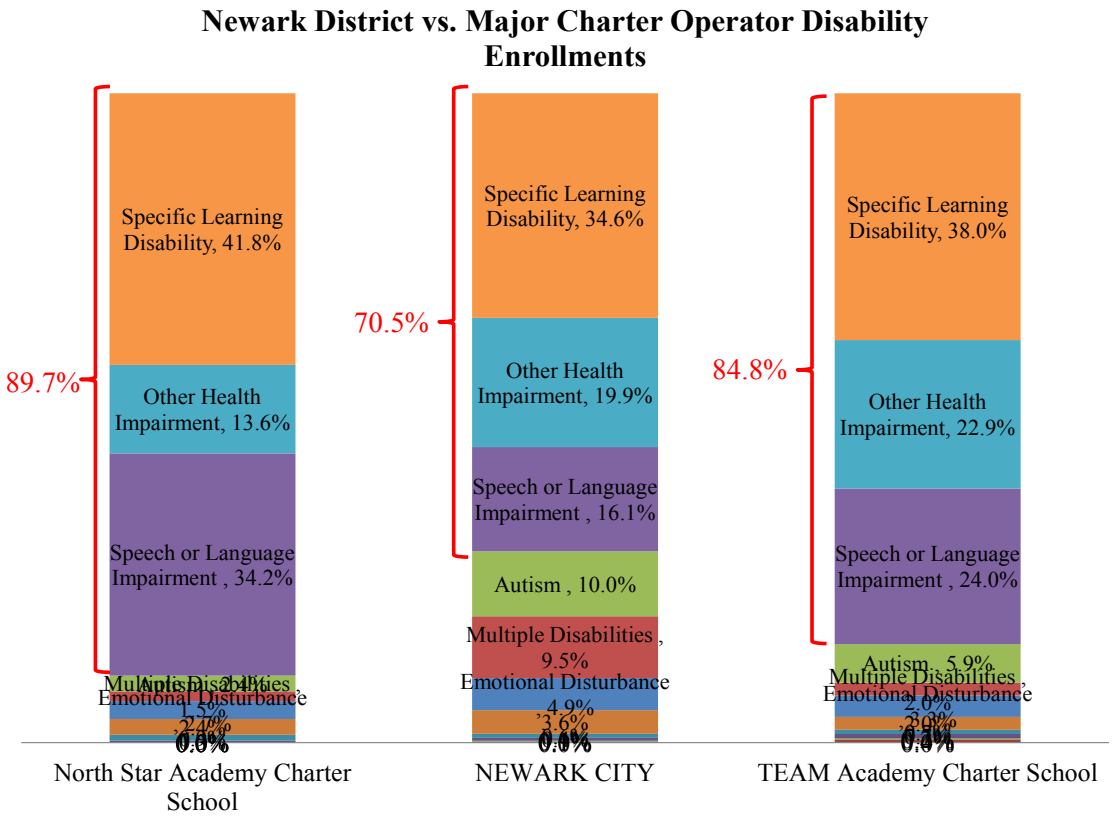


NJDOE Special Education Classification Rates:
http://www.nj.gov/education/specialed/data/2016/LEA_Classificatiom.xlsx

Figure C3 provides a more detailed breakdown, revealing that a very large share of North Star's disability population are children with Speech/Language impairment, and no particular cognitive, behavioral, or other severe impairment which would either divert more substantial shares of resources or directly influence student achievement growth.

Most analyses of Newark district and charter school performance, matching on or controlling for disability status in the aggregate, presume that these children in North Star are equivalent to children with far more severe disabilities in NPS. Some studies specifically find that children with disabilities in charter schools show greater gains than children with disabilities in district schools.[17] In this case (and most other contexts we've studied), such a finding – applying a single measure of "disability" – would be spurious, in that obviously children with only speech language impairment on average would achieve greater growth on standardized assessments than children with multiple and severe learning disabilities.

Figure C3



NJDOE Special Education Classification Rates:
http://www.nj.gov/education/specialed/data/2016/LEA_Classificatiom.xlsx

To summarize, these disability population differences alone, which go unmeasured when using a single "has disability" dummy variable, affect:

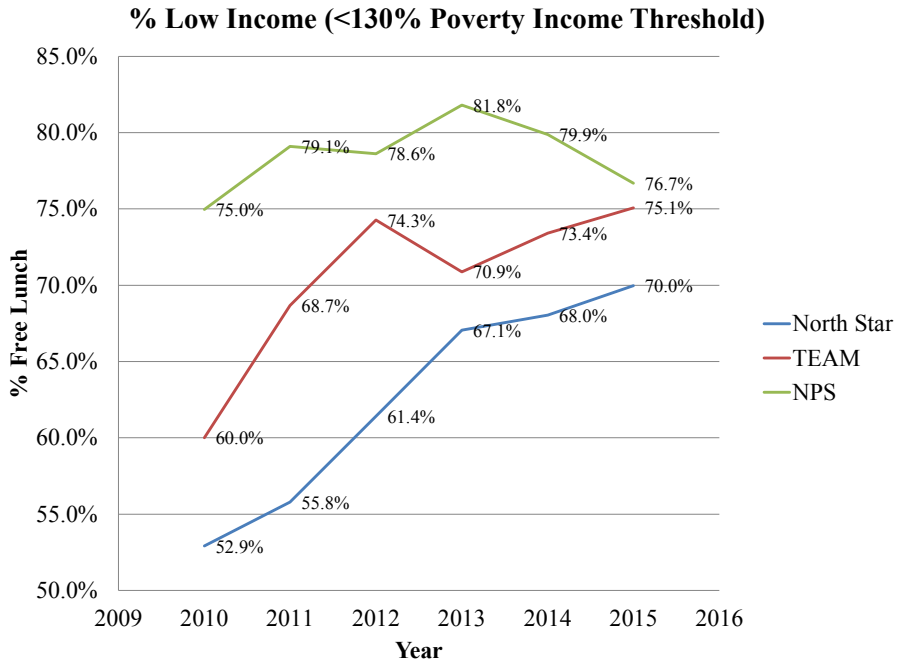
- relative growth between charter and district school students,
- the nature of peer groups (proportions of marginal vs. more severe disability students integrated into regular classrooms could affect the pace of the curriculum and disruptions in classroom time, which likely would affect growth),
- the extent to which higher need student populations create resource pressures and drive reallocation away from "general education" students.

While on the one hand these population differences raise questions regarding assumptions about the effectiveness of North Star Academy, they also raise questions about the scalability of "North Starring" and its effects on the system as a whole, even if North Star is particularly effective with the students that it does serve and retain. The more non-disabled students a single large district in the city enrolls, the more disabled students the other districts will have to serve.

Low Income Concentrations

During the "reform" period under study, substantive differences in the shares of children qualified for "free" lunch existed. These gaps have been closing in recent years; however, North Star continues to serve a smaller share of children who fall below the 130% income threshold for poverty than either TEAM or NPS.

Figure C4



NJDOE Staffing files, 2009-2016.

The Chin et al. study compares students only on the basis of "free + reduced" priced lunch. Single dummy variables on free and reduced-price lunch are relatively meaningless in a context where nearly all children fall below the higher threshold (less than 185 percent of the income poverty level). In fact, those qualified for reduced price lunch are among the more relatively "advantaged" students in the district and schools with higher shares of those students tend to have higher average scale scores.

Table C1 shows the correlations between percent free lunch, percent reduced-price lunch, percent free and reduced-price lunch, and growth and scale score outcome measures across Newark Schools, including district and charter schools. To summarize:

- Percent free lunch has a small negative correlation with growth percentiles and a large negative correlation with scale scores.
- Percent reduced lunch is *positively* correlated with growth and strongly *positively* correlated with scale scores.

- Percent free and reduced-priced lunch is only modestly negatively correlated with scale scores.

This is because those students from families between the 130% and 185% income threshold for poverty happen to be the more "advantaged" students in this high-poverty, urban setting. That is, at the school level, percent free and reduced-priced lunch tells us little about the "risk" of low performance largely because nearly all children in Newark fall below the 185% income threshold for poverty. In addition, it is likely that a substantial number of those who are not identified as qualifying for either in fact do qualify, yet are not listed as such because their families did not apply.

By extension, using a single dummy indicator as a covariate in student (or school) level analysis that assumes nearly all Newark students are socioeconomically identical to one another will lead to specious findings. Because shares of lower income children vary systematically by sector – between NPS and charters – those conclusions will be biased in favor of charters generally, and North Star specifically. While North Star has shown impressive unconditional growth, it has continued to serve fewer of the poorest children in the city. TEAM also served fewer of the poorest children throughout the period studied.

Table C1
Correlations between Growth, Achievement Level and Low Income Populations in Newark (2016)

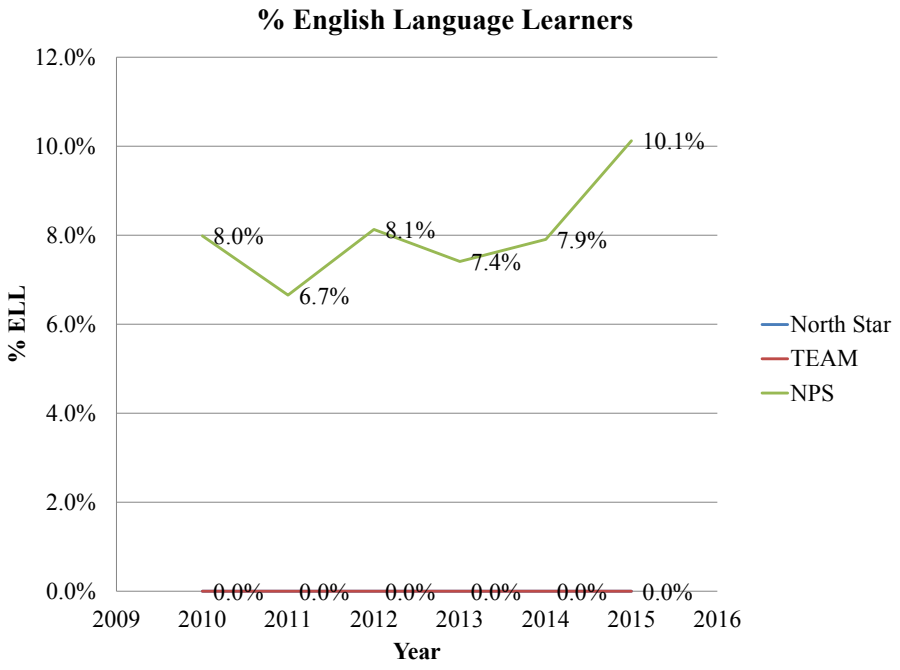
	LAL SGP	Math SGP	PARCC Math 8	PARCC ELA 8	% Free	% Reduced
LAL SGP	1					
Math SGP	0.5807*	1				
PARCC Math 8	0.3758*	0.4686*	1			
PARCC ELA 8	0.4836*	0.4465*	0.9043*	1		
% Free	-0.0984	-0.0734	**-0.3890***	**-0.5052***	1	
% Reduced	**0.3440***	**0.3817***	**0.6602***	**0.8062***	-0.1233	1
% Free or Reduced	0.0444	0.0779	-0.1638	-0.223	0.9348*	0.2373

In addition to compromising validity of high versus low value-added findings, the tendency of between-school mobility to sort students by income status raises scalability concerns. Put bluntly: as one school/district in a high-poverty "choice" space serves more of the less-poor (among the poor) students, others must pick up the difference. Concentrating higher-poverty populations in specific schools potentially creates negative peer effects that are not picked up when using test score histories as measures of peer characteristics.

English Language Learners

Figure C5 shows that among the three districts in Newark, only NPS serves any children with limited English language proficiency. As about 10% of the NPS population is LEP/ELL, this, again, raises questions about scalability. The more that charters in the space serve non-LEP/ELL children, the more LEP/ELL children are concentrated in the district schools. As with poverty and disability, it is also desirable to have access to more fine grained data on the level of language proficiency.

Figure C5
There remain large differences in shares of English Language Learners Served

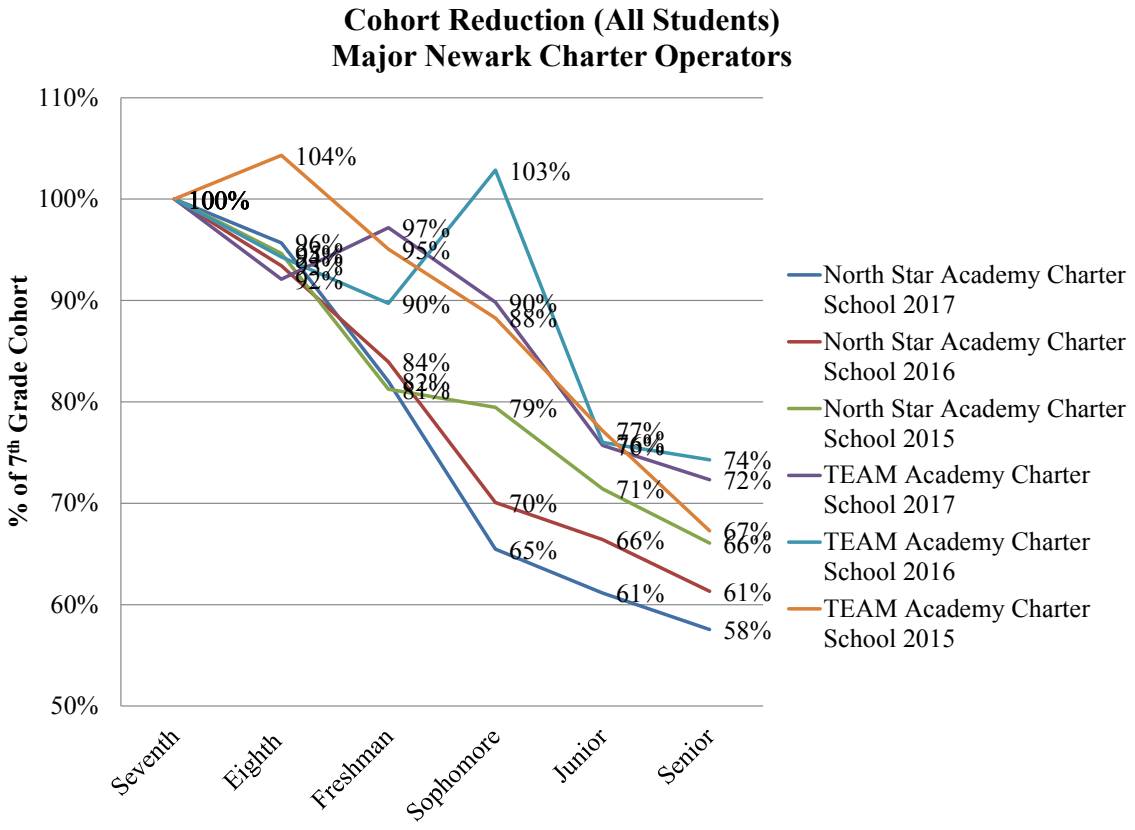


New Jersey Department of Education School Enrollment Files: http://www.nj.gov/education/data/enr/

Cohort Attrition Rates

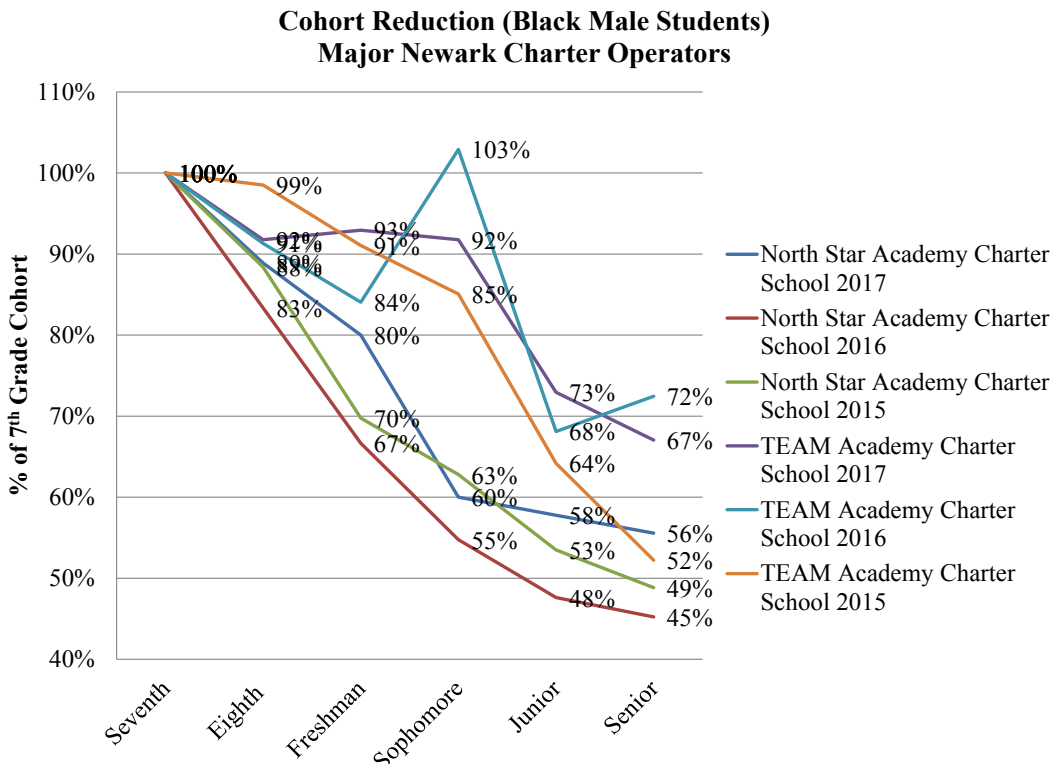
Figure C6 and Figure C7 track cohort attrition rates for three sequential cohorts attending TEAM and North Star. Figure C6 shows the total cohort enrollments and Figure C7 shows the cohort enrollments for black male students. Figure C8 shows the average ratio of the 12th grade enrollment to the 7th grade enrollment of the same cohort of students.

Figure C6
Seventh Grade Cohorts, year after year, are reduced by 25 to 40% as they matriculate to 12th grade



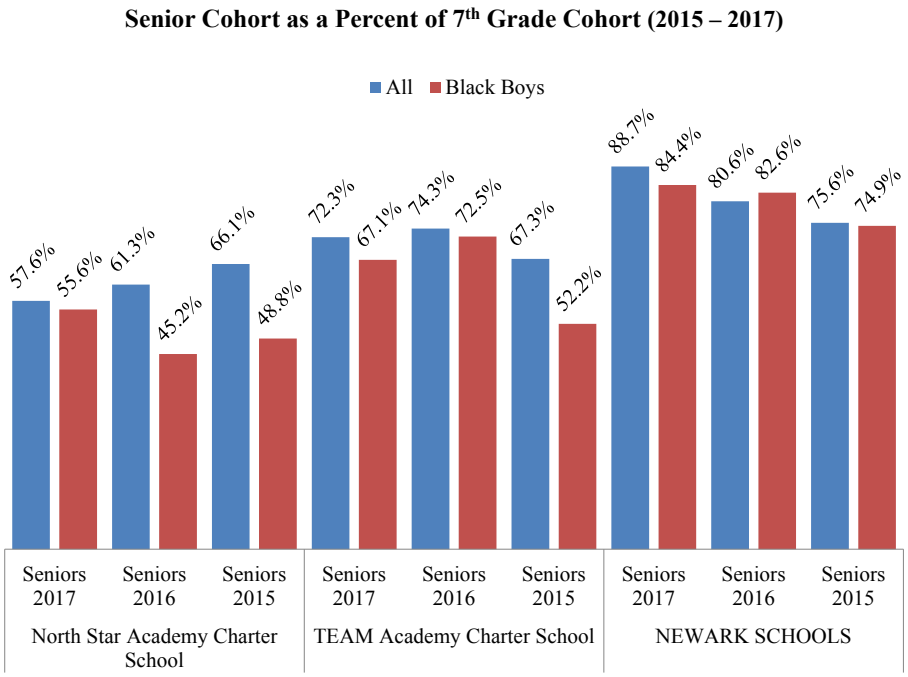
New Jersey Department of Education School Enrollment Files: http://www.nj.gov/education/data/enr/

Figure C7
Seventh Grade Cohorts of Black Boys, year after year, are reduced by 28 to 65% as they matriculate to 12th grade



New Jersey Department of Education School Enrollment Files: http://www.nj.gov/education/data/enr/

Figure C8
Cohort progression rates are much higher for Newark Public Schools than for TEAM and North Star



New Jersey Department of Education School Enrollment Files: http://www.nj.gov/education/data/enr/

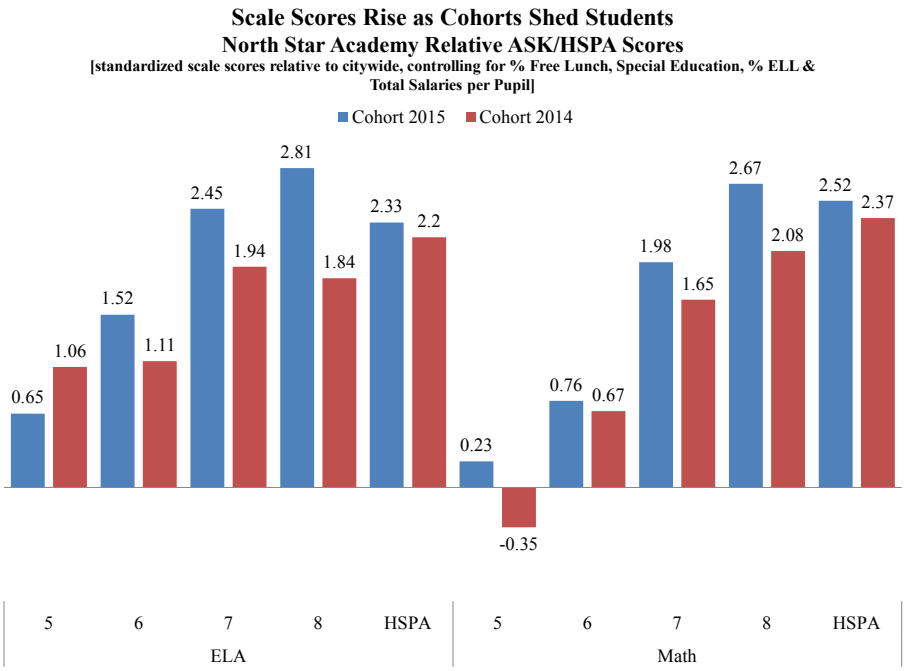
Certainly much can go on between 7th and 12th grade which affects these cohort enrollments. Students can be held back, boosting the prior grade in the subsequent cohort. Cohort reduction might be mitigated by what is called "back-filling" – admitting new students to fill the spaces of students who leave. Also, after 8th grade, some students may choose to leave for other schools, including selective magnet or private schools.

However, if a cohort by 12th grade is substantively smaller than it was in 7th grade, the most likely explanation is that students have left. This cohort attrition might include those who were pushed out and/or counseled out due to behavior or low academic performance, as well as those leaving for private and magnet schools. If the former is true (weaker and "problem" students leaving), we would expect cohort test scores to go up. If, however, the latter is true (students qualified for selective schools leaving), we might expect cohort test scores to go down. Figure C9 addresses this issue.

These figures show that both North Star and TEAM have significant cohort reduction between 7th and 12th grade for all students and even more so for black boys. Senior cohorts of black boys in North Star are half or fewer than the 7th grade cohort.

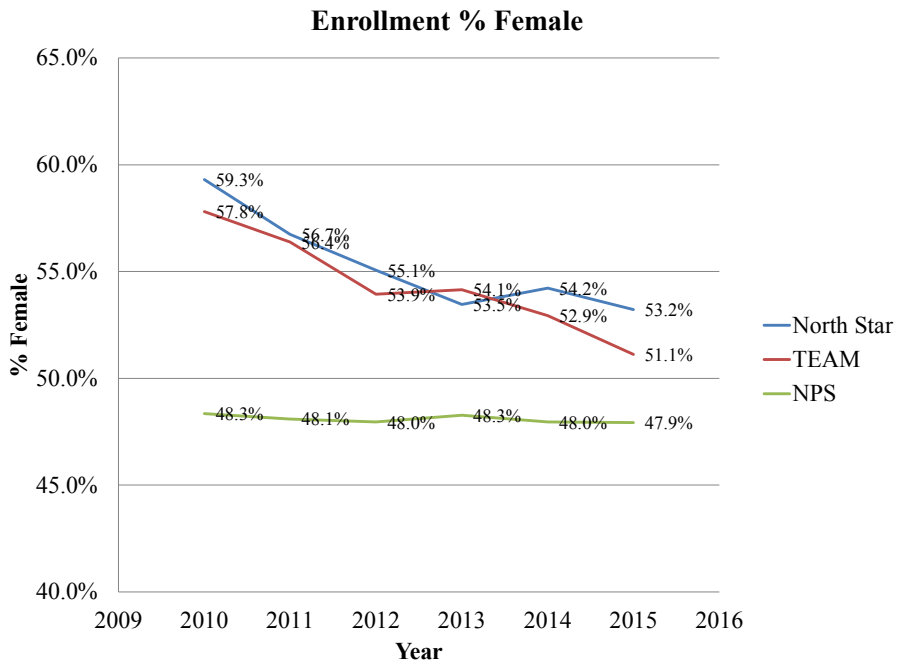
Figure C10 shows that, perhaps in part due to the attrition of black boys over time, these schools also tend to be majority female. As a result, Newark district schools are majority male.

Figure C9
Scale Scores of Cohorts through Progression/Attrition



New Jersey Department of Education School Enrollment Files:
http://www.nj.gov/education/data/enr/
Statewide Assessment Reports:
http://www.state.nj.us/education/schools/achievement/index.html

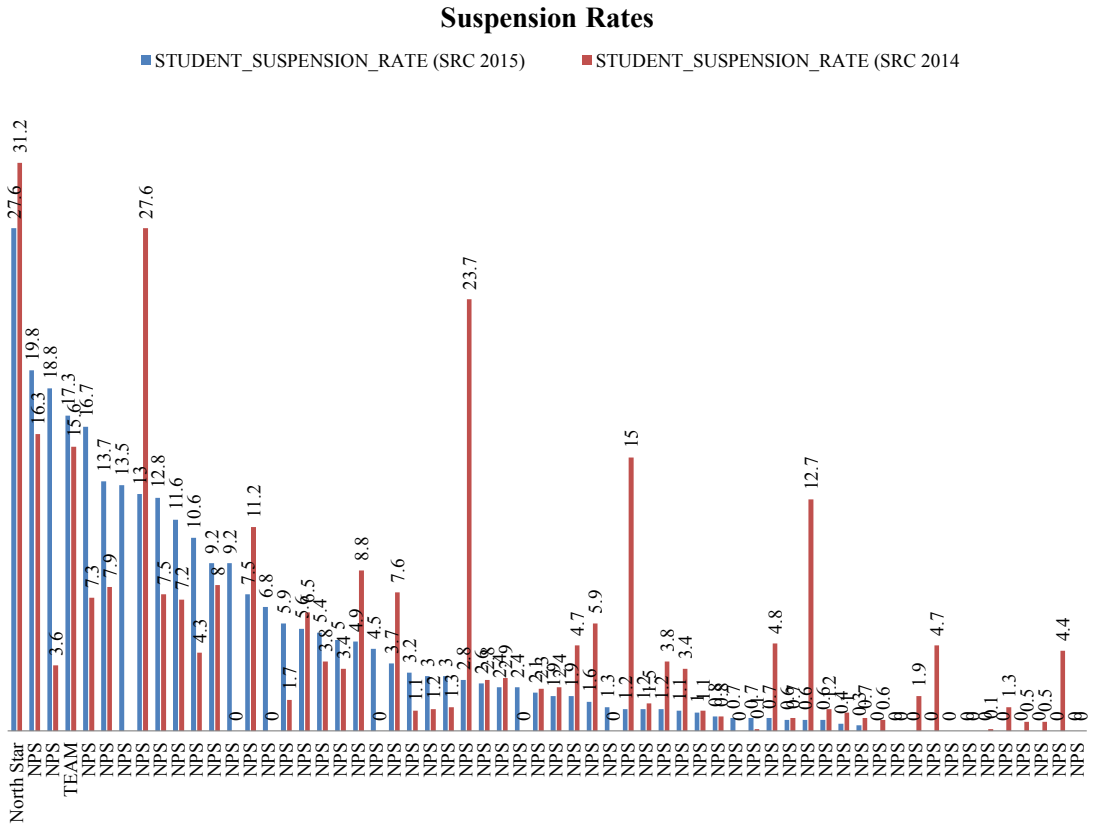
Figure C10
Large charter schools continue to serve predominantly female populations, perhaps as a result of shedding black male students.



New Jersey Department of Education School Enrollment Files: http://www.nj.gov/education/data/enr/

Finally, along with very high attrition rates for black boys, North Star and TEAM Academy continue to have very high student suspension rates. As Figure C11 shows, North Star suspends 30 percent of students year after year.

Figure C11
North Star and TEAM Academy continue to have among the highest suspension rates in the city of Newark (sorted by 2015 rate)



New Jersey Department of Education School Report Cards/ School Climate:
https://rc.doe.state.nj.us/ReportsDatabase.aspx

As a matter of policy preferences for moving forward, these findings raise concerns. Again, the most prominent conclusion from the reports is that citywide gains are achieved by moving more children into high value-added schools, where the largest of those schools – a district within the district - is North Star. North Star's value-added, however, is achieved at least in part (if not in majority) by:

- not serving children with disabilities generally and serving no children with severe disabilities;
- serving very few lower-income children,
- serving no ELL children;
- having very high attrition generally, and 50% or greater attrition of black boys

between 7th and 12th grade; and

- suspending large shares of children year after year.

Having studied these schools year-after-year for nearly a decade, we are confident that these factors taken together are a "feature" and not a bug when it comes to North Star, and remain a feature, though to a lesser extent, in TEAM Academy. These factors are not captured in the reports' analysis. Yet they a) limit the validity of assertions that North Star in particular could be a high value-added school for the general population and b) raise serious concerns regarding policies that would attempt to shift more students to North Star, or schools like it, without first addressing these issues.

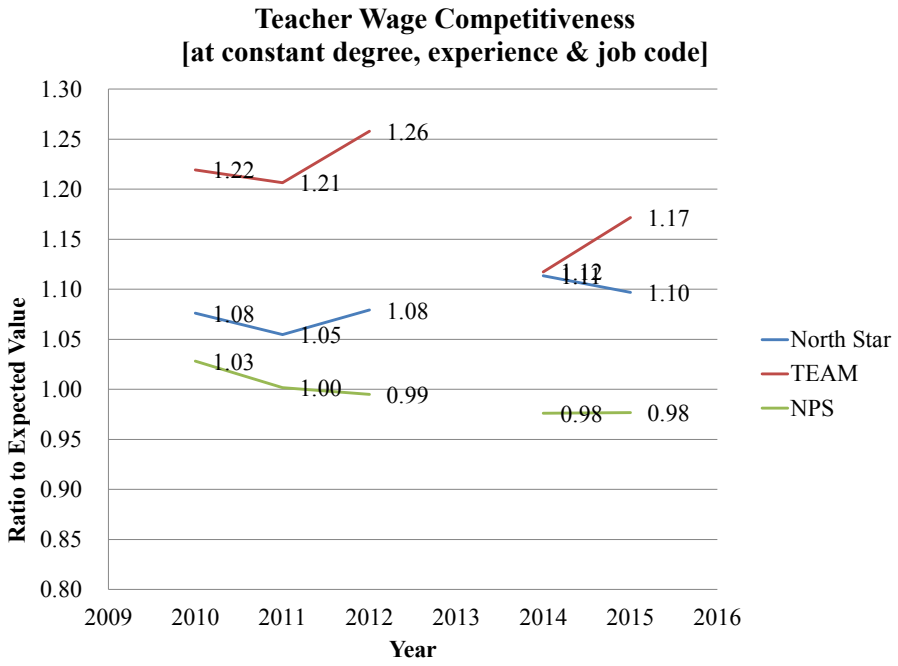
Paying Teachers More to Work More Hours and Days

Here, we address other features of North Star and TEAM as they relate to the host district. These "resource" features may provide more relevant policy insights to the extent that they contribute, in part, to student achievement gains. Resources are legitimately manipulable and scalable features of school systems – at least more so than student sorting by disability and poverty, and selective attrition. Isolating the extent to which these resource factors relate to achievement gains, however, is difficult in the context of these other factors.

Among other things, North Star and TEAM Academy operate longer days (over 8 hours, compared to 6 to 7 for NPS schools, according to state report cards) and longer school years. Figure C12 shows that, on average, teachers in these schools are paid higher wages at similar experience and degree level for this additional time commitment. Teachers in TEAM Academy are paid as much as 20 percent more for their time, compared to teachers with similar characteristics in similar job positions throughout Essex County. Teachers in North Star Academy are paid about 10 percent more. Meanwhile, the relative competitiveness of teacher wages for NPS teachers has slipped below the wage for comparable teachers countywide.

The relevant policy question is: to what extent is this specific investment in teacher wages, for additional time, contributing to the higher value-added at North Star? These differences – time and money – are clearly part of the "treatment" which results from shifting kids from district schools to these two charter operators in particular. Yet this feature of differential treatment between district and charter schools was not addressed in the reports.

Figure C12
Higher pay for longer days and more days



NJDOE Staffing files, 2009-2016.

Relying Heavily on Novice Teachers

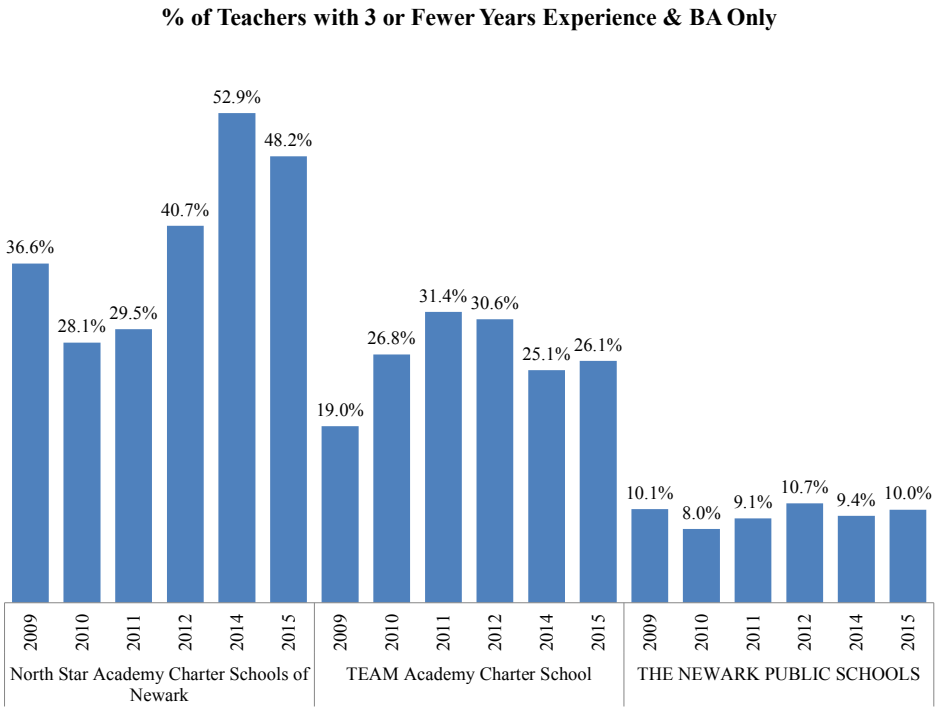
Given the relatively higher wages at TEAM and North Star and the schools' commitment to providing longer days and years, one must question how these schools can keep their ongoing total labor costs under control and sustainable over time. That is, can labor costs be managed in the long run, at even larger scale, while providing 10 to 20 percent compensation increases to support additional contractual time commitments?

Figure C13 provides one answer to how TEAM and North Star have kept their total labor costs in check. These schools maintain staffs with very high shares – up to half – of teachers having three or fewer years of experience. At those experience levels, they are paid more than they would be in the district or elsewhere around the county; however, their average salaries are lower because of their inexperience. TEAM's teaching staff is substantially less novice than is North Star's teaching staff.

One explanation for the large shares of novice teachers in these schools is that they have expanded year after year and have needed new teachers. However, the question remains whether these schools can maintain their approach of longer days and years for higher pay

if these teachers stick around and become more expensive over time. If the model depends on continued turnover to keep spending under control, it may not remain sustainable, especially as it is brought to scale.

Figure C13
Heavy Reliance on Novice Teachers



NJDOE: Staffing files, 2009-2016.

Out-Of-District Peers

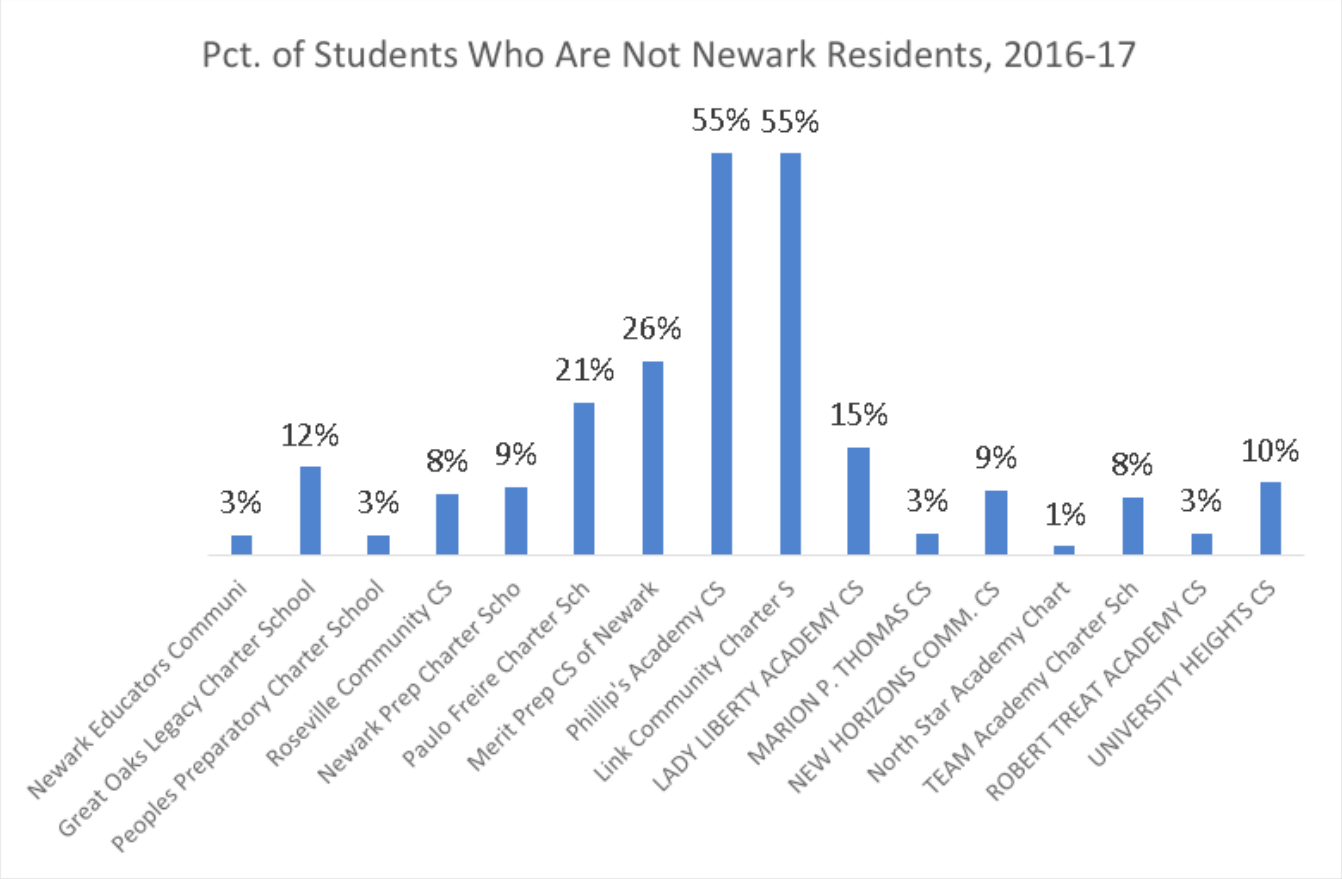
According to state records, a substantial portion of Newark's charter school students are not residents of the district. In New Jersey, charter school funding comes from the district where charter students reside. We use the state's charter aid notices[18] to those districts to calculate the percentages of students who reside outside of the district. In 2016-17, 8 percent of Newark's charter school students are not residents of the city.

Figure C14 shows the percentages of non-resident students by individual charter school. Over half of the student population at two of Newark's charters do not reside in the city. Notably, 8 percent of TEAM/KIPP's students are not Newark residents, while North Star has the highest proportional enrollment of students living in Newark.

It is likely that students who have the ability to travel to another district have unobserved differences in their personal characteristics compared to students who cannot travel. This creates a potential bias in estimates that are derived from comparing non-resident charter

students to resident NPS students.

Figure C14



NJDOE, FY17 Charter School State Aid Notices.

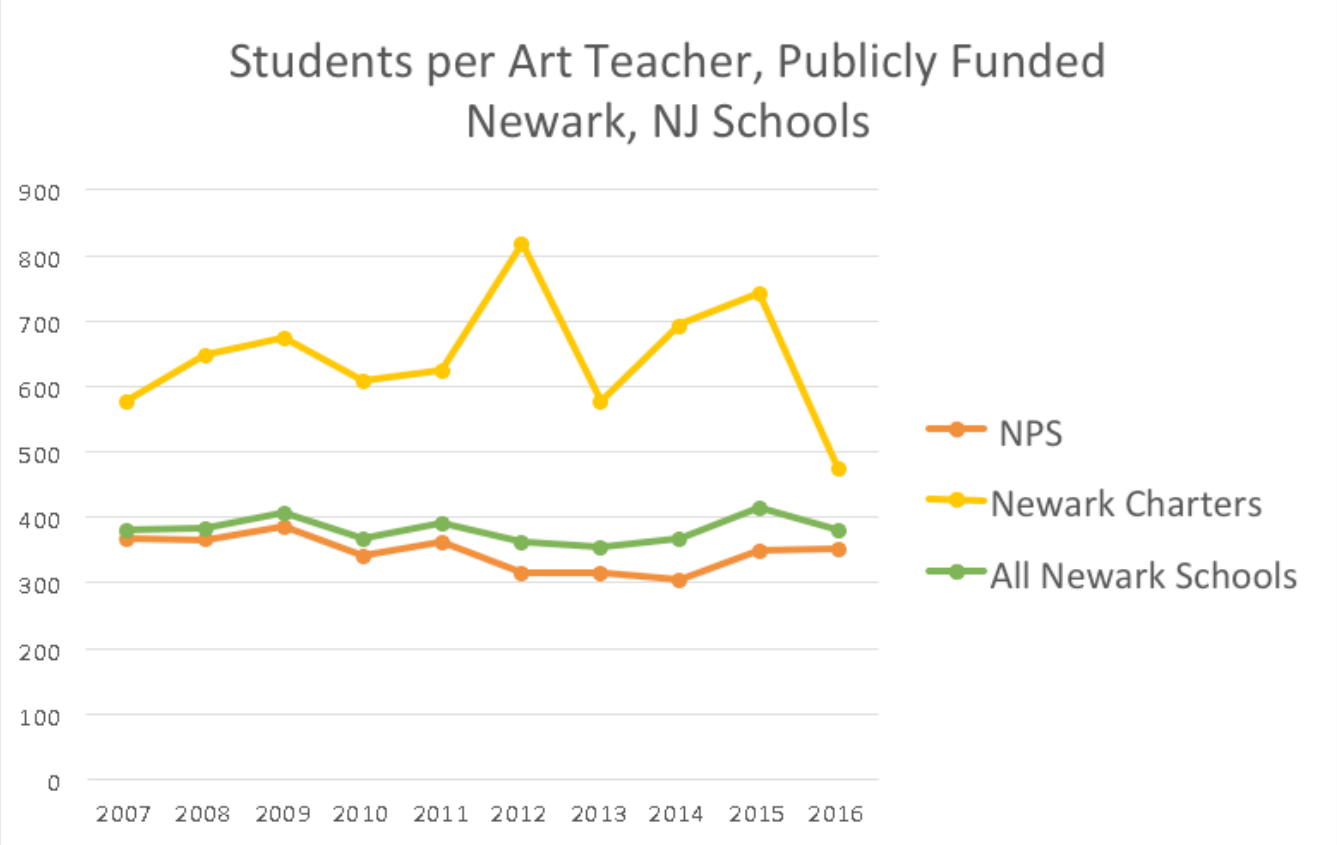
Staff Certifications and Curricular Narrowing

Programs in the arts, physical education, social studies, science, and other "non-tested" subjects require teachers who are certificated in those domains. To the extent that one school has fewer teachers (proportional to student enrollment) with a particular certification than another, we would assume that school offers less extensive programming within that certifications aligned field of study. Put simply: a school with more art teachers per 100 students will likely have more offerings in the arts.

We present here several graphs that show, over a ten-year period, how the Newark charter sector differs from the NPS district in how many teachers in particular subject areas are deployed. Our measure is "student loads": the number of students each teacher certificated in a particular subject would have to teach if the students were all divided evenly among teachers.

Figure C15, for example, shows the student load for art teachers[19] in NPS schools, the charter sector, and all publicly funded Newark schools combined. In every year, art teachers in charter schools have much greater student loads than in NPS. While not definitive proof, this deployment of staff may indicate that charters do not offer coursework in art that is as extensive as NPS schools.

Figure C15

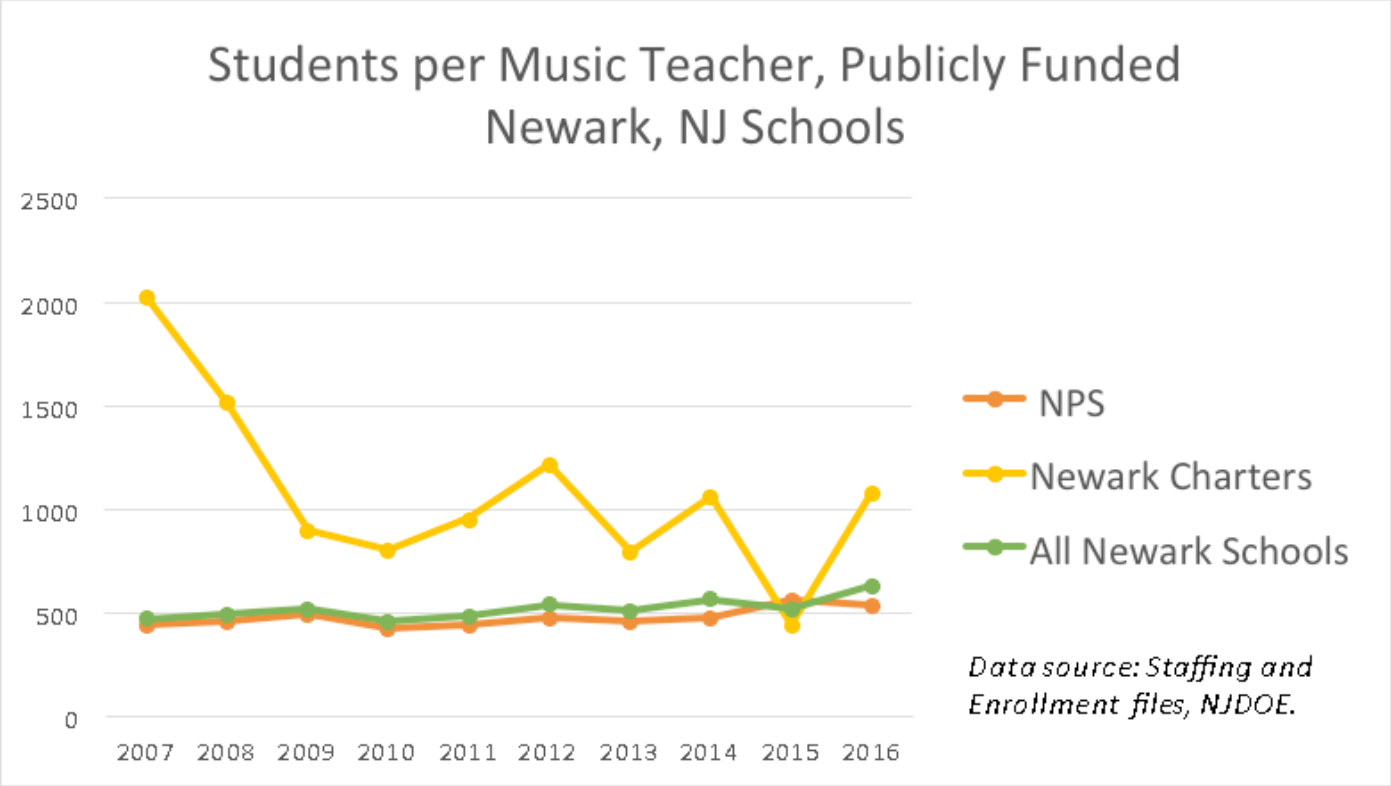


NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/
Staffing files, 2009-2016.

While this data does show significant volatility in the charter schools, the general trend across the past decade has been that Newark charter schools do not have as many staff per student in a variety of non-tested subjects.

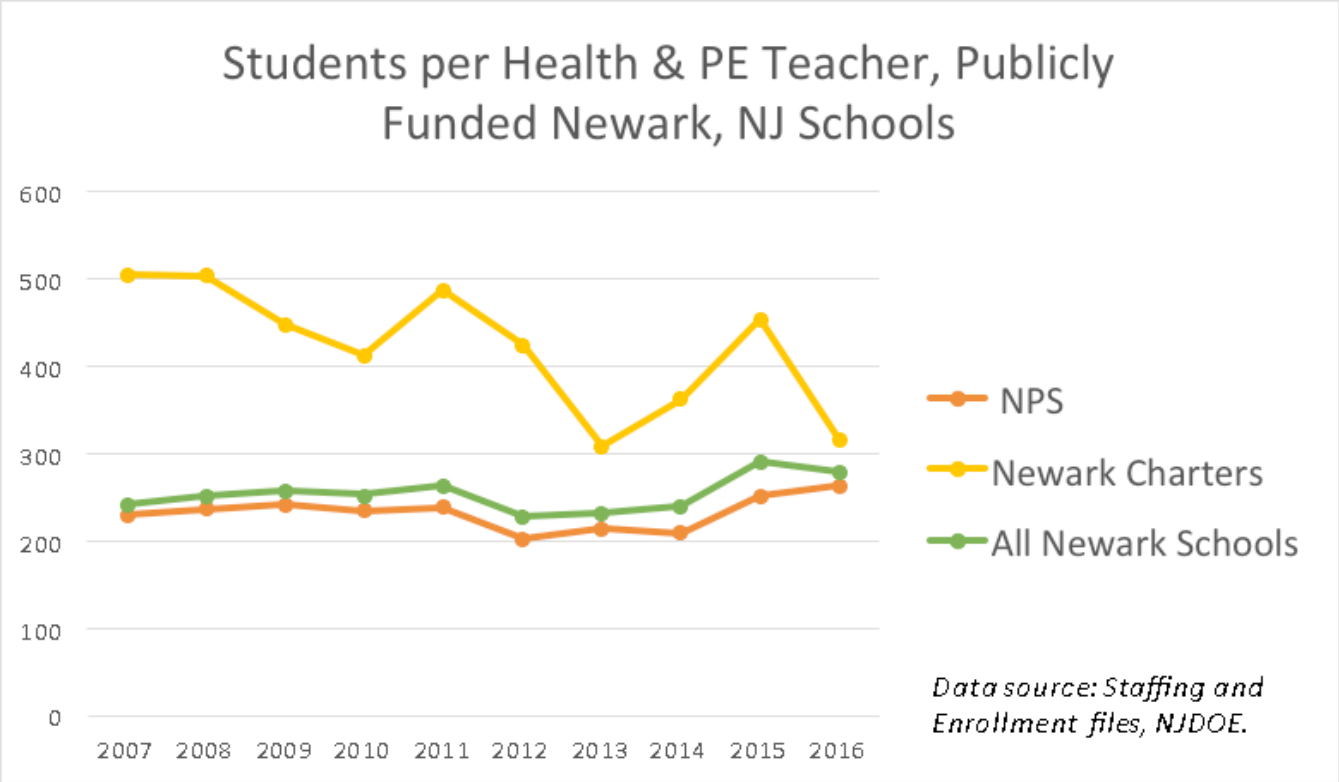
One caution: part of the disparity in staff may be due to differences in grade enrollments. If charters, for example, enroll a smaller proportion of high school students, they may have less need for teachers with social studies certifications. We have begun a preliminary investigation into this possibility. As of now, we do not find that the percentage of Grade 9 to 12 students in a school fully explains the difference between NPS and charter schools. Further analysis, however, may yield different results.

Figure C16



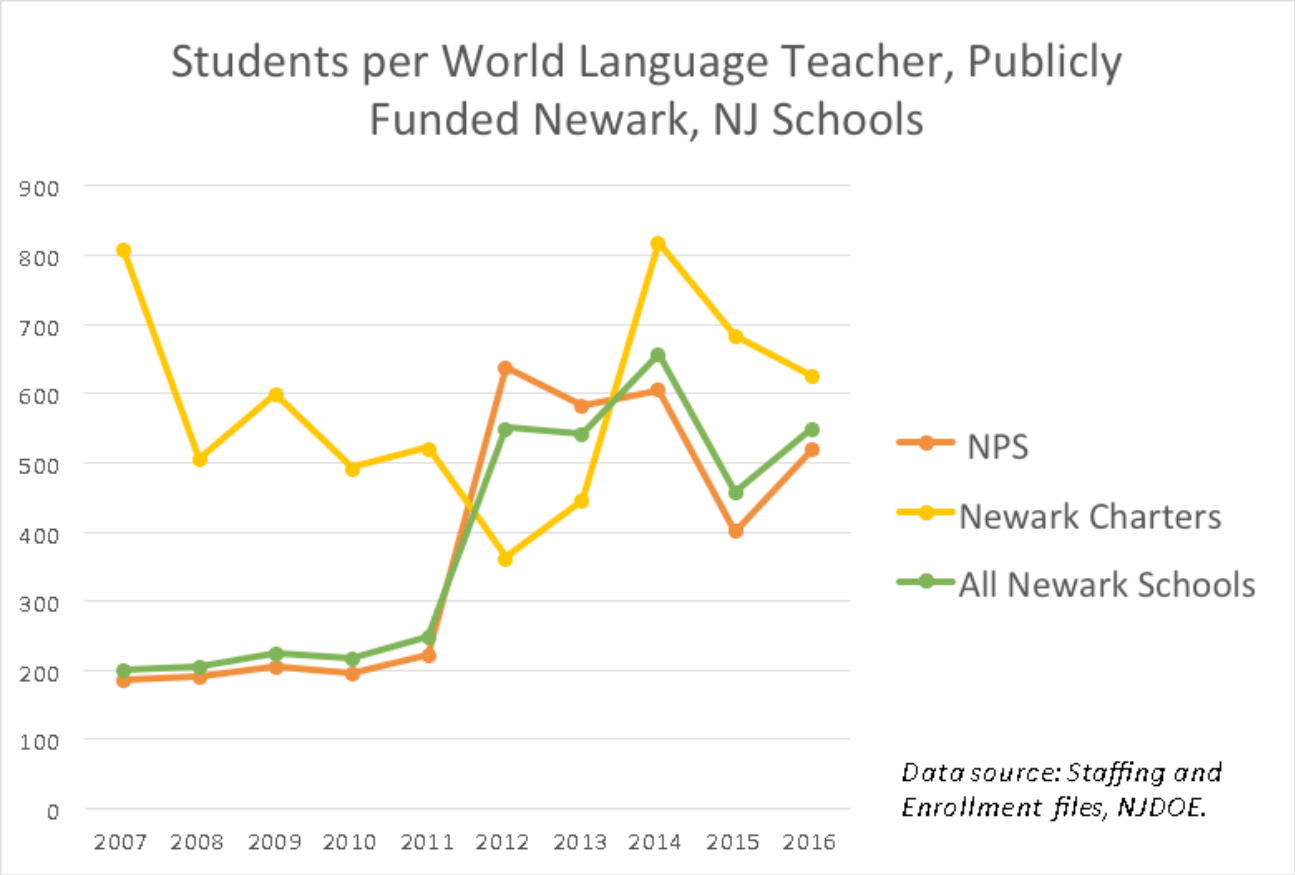
NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/
Staffing files, 2009-2016.

Figure C17



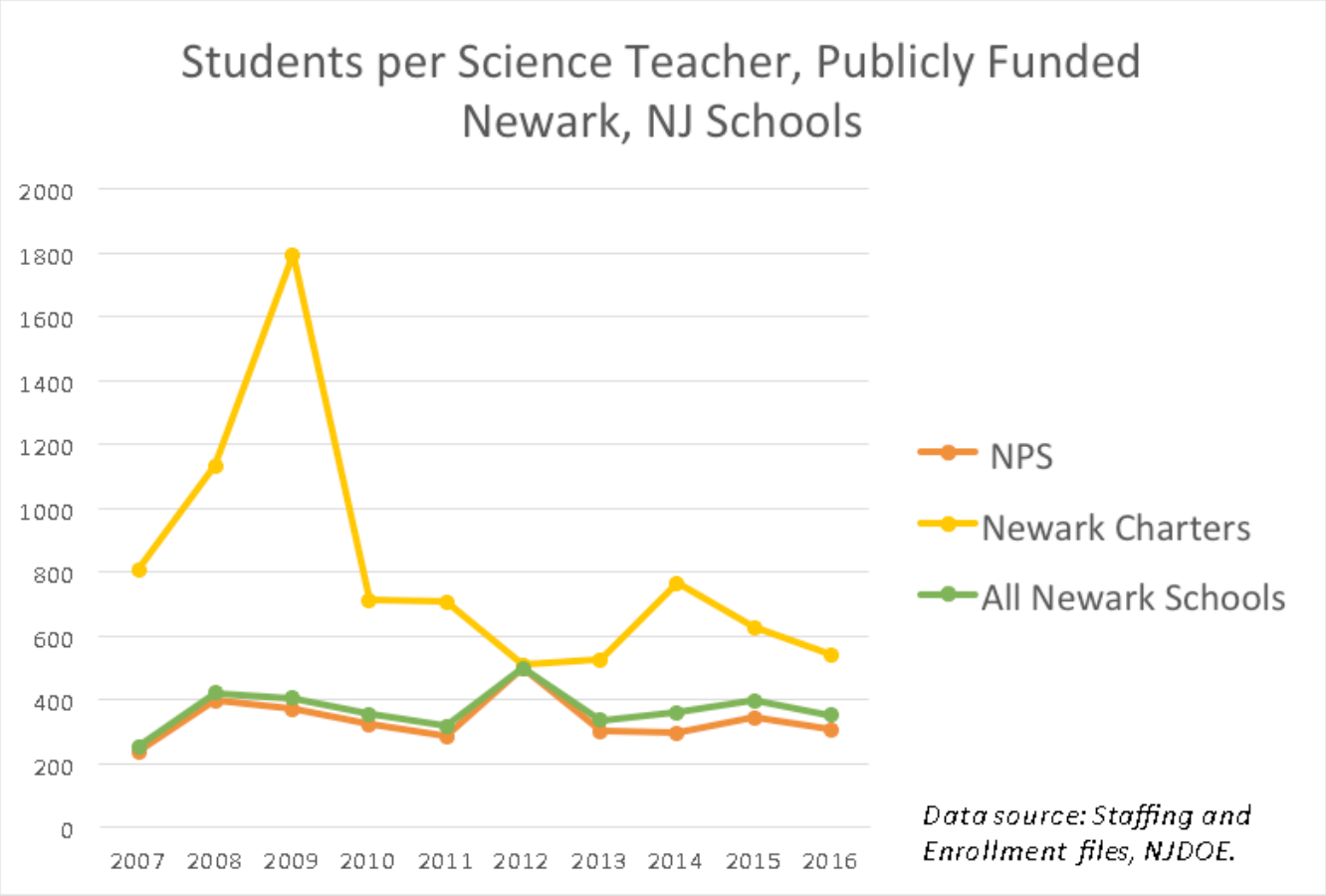
NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/
Staffing files, 2009-2016.

Figure C18



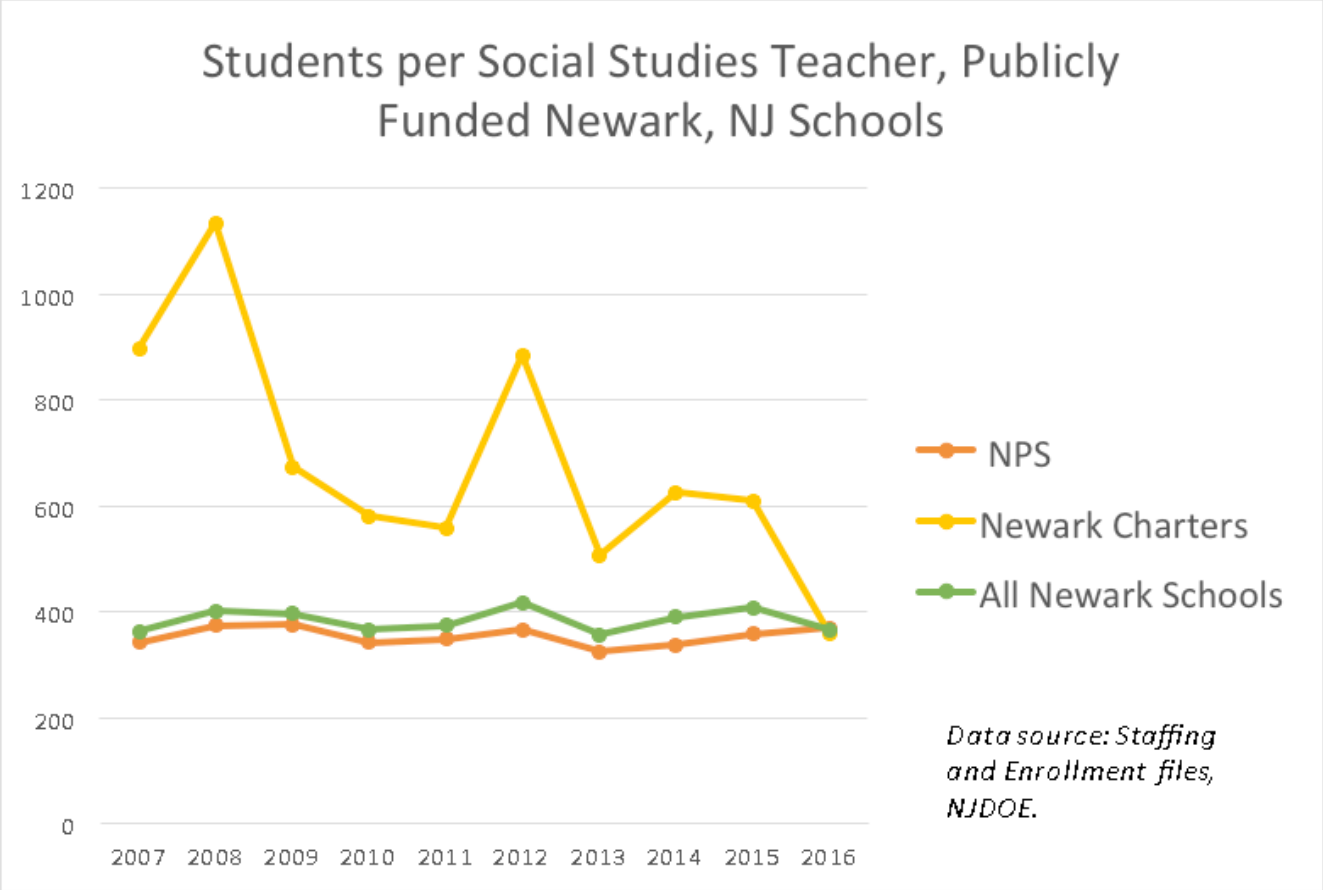
NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/
Staffing files, 2009-2016.

Figure C19



NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/ Staffing files, 2009-2016.

Figure C20



NJDOE: School Enrollment Files: http://www.nj.gov/education/data/enr/
Staffing files, 2009-2016.

Appendix D: Previous Research on Newark Schools

Baker, B. (2016, November 30). *Exploring the consequences of charter school expansion in US cities*. Washington, DC: Economic Policy Institute. Retrieved November 26, 2017, from http://www.epi.org/publication/exploring-the-consequences-of-charter-school-expansion-in-u-s-cities/

Baker, B.D., & Weber, M.A. (2016). *On the Relative Efficiency of New Jersey Public School Districts*. New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2016/08/02/on-the-relative-efficiency-of-new- jersey-public-school-districts/

Weber, M.A. (2016). *New Jersey Charter Schools: A Data-Driven View, Part II - Finances and Staffing*. Princeton, NJ: Daniel Tanner Foundation. Retrieved November 26, 2017, from http://www.saveourschoolsnj.org/nj-charter-school-data/

Baker, B. & Miron, G. (2015). *The Business of Charter Schooling: Understanding the Policies that Charter Operators Use for Financial Benefit*. Boulder, CO: National Education Policy Center. Retrieved November 26, 2017, from http://nepc.colorado.edu/publication/charter-revenue

Baker, B.D., & Weber, M.A. (2015). *Research Note: On Average, Are Children in Newark Doing Better?* New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2015/11/15/research-note-on-average-are- children-in-newark-doing-better/

Weber, M.A. (2015). *The (Mis-)Use of Data in Dale Russakoff's "The Prize."* New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2015/09/21/the-mis-use-of-data-in-dale- russakoffs-the-prize/

Weber, M.A. (2015). *One Newark: Choosing "Great" Schools, or Merely Segregated Ones?* New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2015/04/24/one-newark-choosing-great- schools-or-merely-segregated-ones/

Weber, M.A. (2015). Empirical Critique of "One Newark": First Year Update. Presented to the NJ Legislature's Joint Committee on the Public Schools, March 11, 2015.

NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2015/03/12/empirical-critique-of-one-newark-first-year-update/

Weber, M.A. (2014). *Buyer Beware: One Newark and the Market For Lemons*. New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/05/23/buyer-beware-one-newark-and-the- market-for-lemons/

Baker, B.D., & Weber, M.A. (2014). *An Empirical Critique of "One Newark."* New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/01/24/new-report-an-empirical-critique- of-one-newark/

Baker, B.D., & Weber, M.A. (2014). *A Response to "Correcting the Facts about the One Newark Plan: A Strategic Approach To 100 Excellent Schools."* New Brunswick, NJ: NJ Education Policy Forum. Retrieved November 26, 2017, from https://njedpolicy.wordpress.com/2014/03/24/a-response-to-correcting-the-facts- about-the-one-newark-plan-a-strategic-approach-to-100-excellent-schools/

Weber, M.A. (2014). *New Jersey Charter Schools: A Data-Driven View, Part I - Enrollments and Student Demographics*. Princeton, NJ: Daniel Tanner Foundation. Retrieved November 26, 2017, from http://www.saveourschoolsnj.org/nj-charter-school-data/

Endnotes

1 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, Figure 1.

2 Education Law Center (n.d.) The history of Abbott v. Burke. Newark, NJ: Education Law Center. Retrieved November 26, 2017, from http://www.edlawcenter.org/cases/abbott-v-burke/abbott-history.html

3 Baker, B. (2016). *Exploring the consequences of charter school expansion in US cities*. Washington, DC: Economic Policy Institute. Retrieved November 26, 2017, from http://www.epi.org/publication/exploring-the-consequences-of-charter-school-expansion-in-u-s-cities/

4 Baker, B.D. (2011, October 22), Thoughts on improving SFRA [Blog post]. Retrieved November 26, 2017, from https://schoolfinance101.wordpress.com/2011/10/22/thoughts-on-improving-the-school-funding-reform-act-sfra-in-nj/

5 Baker, B.D. (2014). Evaluating the recession's impact on state school finance systems. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, *22*.

6 Baker, B.D. (2011, May 24) Demystifying Today's Abbott Decision [Blog post]. Retrieved November 26, 2017, from https://schoolfinance101.wordpress.com/2011/05/24/demystifying-today%e2%80%99s-abbott-decision/

7 Baker, B. (2016, November 30). Exploring the consequences of charter school expansion in US cities. Washington, DC: Economic Policy Institute. Retrieved November 26, 2017, from http://www.epi.org/publication/exploring-the-consequences-of-charter-school-expansion-in-u-s-cities/

8 Rinde, M. (2016, May 17) Explainer: Getting the facts on funding for NJ's charter schools. Retrieved November 26, 2017, from http://www.njspotlight.com/stories/16/05/16/explainer-getting-the-facts-on-funding-for-nj-s-charter-schools/

9 We note here that Camden moved to universal free lunch enrollment in September of 2014; see: http://www.camden.k12.nj.us/UserFiles/Servers/Server_340793/File/Migrate/Divisions/Division_of_E/Press_Room/News_Archives/20142015/14_Camden_schools_to_provide_free_breakfast_and_lunch_for_all_District_students.pdf As a consequence, the publicly reported FRPL percentage has decreased in recent years, as families have no need to fill out forms. There is no indication in the reports that the VAMs used to measure achievement growth accounted for this change.

10 Baker, B.D., Taylor, L., Levin, J., Chambers, J., & Blankenship, C. (2013). Adjusted Poverty Measures and the Distribution of Title I Aid: Does Title I Really Make the Rich States Richer?. *Education*, *8*(3), 394-417.

11 Chin, M., Kane, T.J., Kozakowski, W., Schueler, B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 40.

12 The NBER report explains in a footnote: "…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." (Chin, M., Kane, T.J., Kozakowski, W., Schueler,

B.E., & Staiger, D.O. (2017). *School District Reform in Newark: Within- and Between-School Changes in Achievement Growth* (Working Paper No. 23922). National Bureau of Economic Research. Retrieved November 26, 2017, from http://www.nber.org/papers/w23922, p. 43

13 Baker, B.D. (2003). State policy influences on the internal allocation of school district resources: Evidence from the common core of data. *Journal of Education Finance*, *29*(1), 1-24.

14 http://www.nj.gov/education/data/enr/

15 See, for example: Kingsland, N. (2017). Could Newark have achieved more? [Blog post]. Retrieved November 26, 2017, from https://relinquishment.org/2017/10/23/could-newark-have-achieved-more/

16 Winters, M.A. (2013). Why the gap? Special education and New York City charter schools. *Manhattan Institute for Policy Research and Center for Reinventing Public Education*.

Winters, M.A., Carpenter, D.M., & Clayton, G. (2017). Does Attending a Charter School Reduce the Likelihood of Being Placed Into Special Education? Evidence From Denver, Colorado. *Educational Evaluation and Policy Analysis*, 0162373717690830.

17 Center for Research on Education Outcomes (2015). Urban charter school study report on 41 regions. Retrieved November 26, 2017, from https://urbancharters.stanford.edu/download/Urban%20Charter%20School%20Study%20Report%20on%2041%20Regions.pdf This study is cited by the reports to assert that "Newark is home to one of the most effective charter sectors in the nation in terms of student growth on standardized exams." (p. 19)

18 We thank Dr. Julia Sass Rubin of the Bloustein School of Planning and Public Policy, Rutgers, the State University of New Jersey, for the data.

19 For each of the categories given, we consolidate job codes into larger categories. For example: "art teachers" include photography, ceramics, theatre/stage, dance, etc. We use the categories provided by NJDOE for guidance.