NEPC Review: Charter Schools in Newark: The Effect on Student Test Scores (Manhattan Institute, January 2020)

Reviewed by:

Mark Weber
New Jersey Policy Perspective and Rutgers University

March 2020

National Education Policy Center

School of Education
University of Colorado Boulder
nepc.colorado.edu

Acknowledgements

NEPC Staff

Kevin Welner
NEPC Director

William Mathis
Managing Director

Alex Molnar
Publications Director

Suggested Citation: Weber, M. (2020). *NEPC Review: "Charter Schools in Newark: The Effect on Student Test Scores."* Boulder, CO: National Education Policy Center. Retrieved [date] from http://nepc.colorado.edu/thinktank/newark-charters

Funding: This review was made possible in part by funding from the Great Lakes Center for Educational Research and Practice.

NEPC Review: Charter Schools in Newark: The Effect on Student Test Scores (Manhattan Institute, January 2020)

Reviewed by:

Mark Weber
New Jersey Policy Perspective and Rutgers University

March 2020

Executive Summary

Charter Schools in Newark: The Effect on Student Test Scores, published by the Manhattan Institute, attempts to estimate the effects of charter school enrollment on student test scores. The report exploits a random component of Newark's school enrollment system to isolate the effect of charter schools on outcomes. This report's method credibly creates a natural experiment that compares outcomes between students who were offered charter seats and those who were not. However, the students who apply to charter schools differ significantly from the greater Newark student population. Specifically, students who are offered seats in charter schools are less likely to be English language learners or have special education needs. But the report never addresses this core issue of external validity. The report also fails to note that teachers in the studied charter schools are less likely to remain in their jobs more than a few years, with the result that the schools employ relatively few experienced teachers. This allows the schools to pay their teachers more than public district school teachers with the same experience and to provide longer school days at a lower cost, likely positively impacting test scores. The report fails to address a critical policy question: Can these charter schools' financial models be sustained as they expand and enroll a different student population? While this report confirms previous work showing a positive charter effect on test scores in Newark, it fails to address these and other critical issues stakeholders must understand to formulate good charter school policy in New Jersey and elsewhere. Accordingly, while the report has credible internal validity, it has only limited policy use.

NEPC
NATIONAL EDUCATION POLICY CENTER

NEPC Review: Charter Schools in Newark: The Effect on Student Test Scores (Manhattan Institute, January 2020)

Reviewed by:

Mark Weber
New Jersey Policy Perspective and Rutgers University

March 2020

I. Introduction

Newark's charter schools have been a subject of intense scrutiny over the past several years. A high-profile gift from Facebook CEO Mark Zuckerberg in 2010[1] and the presidential candidacy of former Newark mayor Cory Booker kept the spotlight on Newark's charter schools over the last decade.

In *Charter Schools in Newark: The Effect on Student Test Scores*,[2] Marcus Winters, senior fellow at the Manhattan Institute (which published the report), attempts to exploit the random component of Newark's school enrollment system to create a natural experiment that estimates the effectiveness of the city's charter schools.

II. Findings and Conclusions of the Report

The report finds that students who enrolled in a Newark charter school that participates in Newark Enrolls (the city's common school enrollment system) saw an increase of 0.263 standard deviations (SD) in math and 0.246 SD in English Language Arts (ELA) in the first year. The report considers these effects "large."

The report also estimates effects separately for charter schools run by either KIPP-NJ or Uncommon Schools[3], two large charter management organizations (CMOs) with national reach. The effect sizes for these charters are much larger than the other Newark Enrolls charter schools; the report conjectures, based on previous studies,[4] that these CMOs' "no excuses" approach is responsible for the magnitude of the effect.

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

The report places great importance on the use of random assignment to charter schools in its design. Newark Enrolls uses a deferred assignment (DA) system that takes into account student preferences before offering students seats in charter or public schools. While not entirely random, assignment to schools does have a random component, where otherwise similar students' offers are left to chance.

The report replicates a complex model, developed for use in studying Denver's similar school assignment system,[5] that ostensibly isolates the random component of the DA system. This creates an experimental condition that controls for both observed and unobserved student characteristics. The results from such experiments can be credibly said to be caused by the "treatment" – in this case, charter school enrollment – rather than unobserved differences between charter and public district students, such as parental involvement, academic talent, etc.

IV. The Report's Use of Research Literature

The report cites a substantial body of literature on the effects of charter school enrollment on student outcomes. Several of these studies exploit the random assignment to schools inherent in some urban enrollment systems;[6] these studies clearly serve as models for the report. One feature of this research is its use of a two-stage model (2SLS) to account for the problem of non-compliance with treatment. The two-stage model has become a standard method for dealing with this research problem;[7] however, the method has properties, described below, that may limit the relevance of its estimates in formulating charter school policy.

V. Review of the Report's Methods

There are at least three aspects of this report's methods that policymakers should consider: 1) the validity of the econometric model, 2) the description of the size of the effect, and 3) the definition of the treatment – i.e. "charter school enrollment" – and whether the definition fully describes *how* charter schools realize their effects.

The Econometric Model

This report leverages the part of the Newark enrollment system that randomly assigns some students to a charter school. A core precept of this type of research is that random assignment to treatment controls for unobserved differences in treatment and control groups; in other words, if Newark students are randomly assigned to attend either a charter or district school, we can assume that the two groups of students are equivalent in all unobserved characteristics. This means we can credibly infer that any differences in their outcomes are due to their enrollment in different types of schools.

Yet even if we accept that the model credibly takes advantage of the randomization, there are two major limitations. First: We can only assume the results found in this report are relevant to the population of students who used the DA system and applied to a charter school. This population is not, however, the same as the population of all Newark students.

There are many reasons to believe the subgroup used in this report differs significantly from the entire population of Newark students; in other words, students who apply to charter schools are not the same as students who do not. This is confirmed by the report itself: 11% of the students who were offered seats in charters are classified as special education, while the entire Newark student population rate[8] is 16% (See Figure 7 from the report). Similarly, the entire Newark student population is seven percent Limited English Proficient (LEP), while the rate for those offered a charter seat is only three percent.

Further, there is reason to believe the larger student population differs from the subgroup offered charter seats on unmeasured characteristics, such as parental involvement, student temperament, academic talent, etc. Parents who learn about a charter school's "no excuses" philosophy, for example, may conclude it is not a good fit for their child, and therefore not apply. Thus, the report's *generalizability* is not solved by the random component of assignment to charters.

The second limitation of the report concerns what researchers often refer to as *compliance with treatment*. Even if students are offered a seat at a charter school, they may not accept the offer. Because only the offer to attend a charter is random, the students who do enroll or do not enroll in charters may differ.

This problem has been described in previous charter studies that use random assignment to charters to estimate the effects of charter school enrollment.[9] The proposed solution, used in this report, is to employ a two-stage regression model. The report explains: "*The model essentially divides the estimated effect of assignment to a charter school by the proportion of assigned students who actually enrolled in a charter*" (p. 9). This means the model will increase the estimated effect as fewer students who are offered charter seats actually enroll in the schools.

There is, however, a question of whether the 2SLS model actually produces estimates that are truly policy-relevant. Consider this example:

Imagine two schools, completely alike in all ways: same faculty, same instructional methods, same facilities, etc. Imagine these two schools enroll a student population exactly the same in all observed and unobserved characteristics: same parental involvement, same rate of special education, etc. Both schools' students take the same pretest; they score exactly the same. After a year, they take a post-test; again, they score exactly the same, because both the students and the schools are alike in every way.

Now imagine there is only one difference between the schools: Even though they wind up with exactly the same student population, one school had more applicants who chose not to enroll after they were offered a seat. Again, the final student populations turn out to be the same, as confirmed by their similar pretest and post-test scores.

In a 2SLS model, the schools will have different estimated effects on student test scores. This is because, as described by the report, the estimated effect is divided by the proportion of students who accept an offer. The result is that two schools, with exactly the same effects on the test scores of exactly the same students, are reported as having different effects if they have different acceptance rates.

While there may be a theoretical justification for 2SLS estimates, their usefulness in informing policymaking is questionable. For this reason, research employing 2SLS models always includes estimates from simple regression models as well as 2SLS models; this way, readers can evaluate the relative differences in the estimates. This report, however, does not include these estimates.

Describing the Size of the Effect

Another problem researchers face in reports of this type is interpreting the estimates of the effect sizes in language that lay readers can understand. This report uses a recent working paper by Matthew Kraft of Brown University[10] to justify denoting these effects as "large." This classification, however, fails to account for one important contextual factor: How does the effect compare to the stated goal of those who advocate for the policy under study?

Both North Star/Uncommon and TEAM/KIPP-NJ have repeatedly compared their students' test scores to those of students in New Jersey's more affluent suburban districts and to state averages.[11] In writing about Uncommon's charter schools, Doug Lemov, Managing Director of Uncommon, asserts: "*These outstanding teachers routinely do what a thousand hand-wringing social programs have found impossible:* ***close the achievement gap between rich and poor****, transform students at risk of failure into achievers and believers, and rewrite the equation of opportunity.*"[12]

Kraft notes that how a standard deviation is calculated – specifically, what data is used – can affect the calculation of an effect size. Considering the report states the data used is for Newark students, a reader would likely assume the effect is limited to the variation in scores across Newark. This makes it difficult for the reader to evaluate whether the effect size is "large" in the context of evaluating whether Newark's charter schools "*close the achievement gap between rich and poor*."

As an example: The 0.263 SD effect on math scores is the equivalent of moving students from the 50th to the 60th percentile on test outcomes. But the reader does not know if that move is across the scores for students who were offered charter seats, or all students in Newark, or all students across the entire state. The report, therefore, sheds little light on whether Newark's charter schools are meaningfully meeting the goals they have set for themselves.

Definition of the Treatment

In 2017, Rutgers professor Bruce Baker and I wrote an extensive report on Newark's schools, including its charter schools, as part of a review of a report on student outcomes published

by the Center for Education Policy Research at Harvard University.[13] One of the takeaways from our review is that research that makes simple classifications of schools as "charter" and "district," and then compares the outcomes of students in each, does little to help illuminate *why* each type of school gets the outcomes that they do. The report asserts that TEAM/KIPP-NJ and North Star/Uncommon both use a "no excuses" approach, but gives only a vague definition of what this approach entails.

Data on Newark's schools, however, can go a long way toward explaining the "success" of the Newark's largest charter school operators. Table 1[14] summarizes several important data points.

Table 1

2017-18 (*2018-19)	Avg. Teacher Experience (Years)	Avg. Teacher Salary*	Avg. Teacher Salary, 5 Years Experience*	Teacher 1-Year Retention Rate	Pct. Special Education*	Pct. Limited English Proficient*	Teacher & Support Salaries Per Pupil	Length of School Day (Minutes)
Newark Public Schools	12.5	$76,543	$62,144	82.0%	16.2%	12.5%	$10,891	390
North Star	3.8	$66,530	$67,142	68.0%	8.3%	0.8%	$5,246	510
TEAM-KIPPNJ	6.2	$71,959	$69,553	44.0%	12.7%	1.7%	$6,027	540
Other Newark Charters	5.2	$61,146	$61,189	69.0%	10.0%	1.6%	$7,630	463

- Newark charter schools have teaching staffs that have much less experience than NPS staff. Consequently, the average charter teacher salary is significantly less than the average for NPS teachers.
- But a comparison of salaries for teachers with only five years of experience shows that teachers in North Star/Uncommon and TEAM/KIPP-NJ schools make considerably more than NPS teachers with the same experience.
- These CMOs are able to keep their salaries for relatively inexperienced teachers competitive, while simultaneously keeping their average salaries low, by churning their staff. TEAM/KIPP-NJ, for example, has a one-year teacher retention rate of only 44%.
- The charter sector enrolls a different student population than NPS – a population that has less costly learning needs. NPS has proportionally more students classified with a special need, and far more students who are Limited English Proficient, than the charter schools.[15]
- NPS students, consequently, need more staff certified in delivering special education

and English as a second language instruction and support. This, combined with the more experienced staff, explains the higher staff costs per pupil at NPS.

- Because the charter school teacher salaries are higher for less experienced staff, these schools can pay their staffs to work a longer school day; North Star/Uncommon's day is two hours longer than NPS's, and TEAM/KIPP-NJ's day is even longer. This extra time likely explains at least some of the learning gains charters show over NPS.

The relevant policy question, given this data, is whether such a system can be brought to a larger scale. If the Newark charter sector continues to expand, it must, under its current staffing and spending model, recruit even more teachers who are willing to work longer hours for more pay, but not to stay on staff for as long as NPS teachers do. It must also continue to enroll students with, on average, fewer higher-cost education needs. Whether this is desirable, or even possible, is an open question.

There are other factors that help explain the effects realized by Newark's charter schools. As Baker and I wrote, the leaders of Newark's biggest charter schools make no secret of their emphasis on test preparation. Test-prep pedagogy can help increase test scores;[16] however, it can also narrow the curriculum, denying students a broad education including content from many domains of learning.

Figure 1

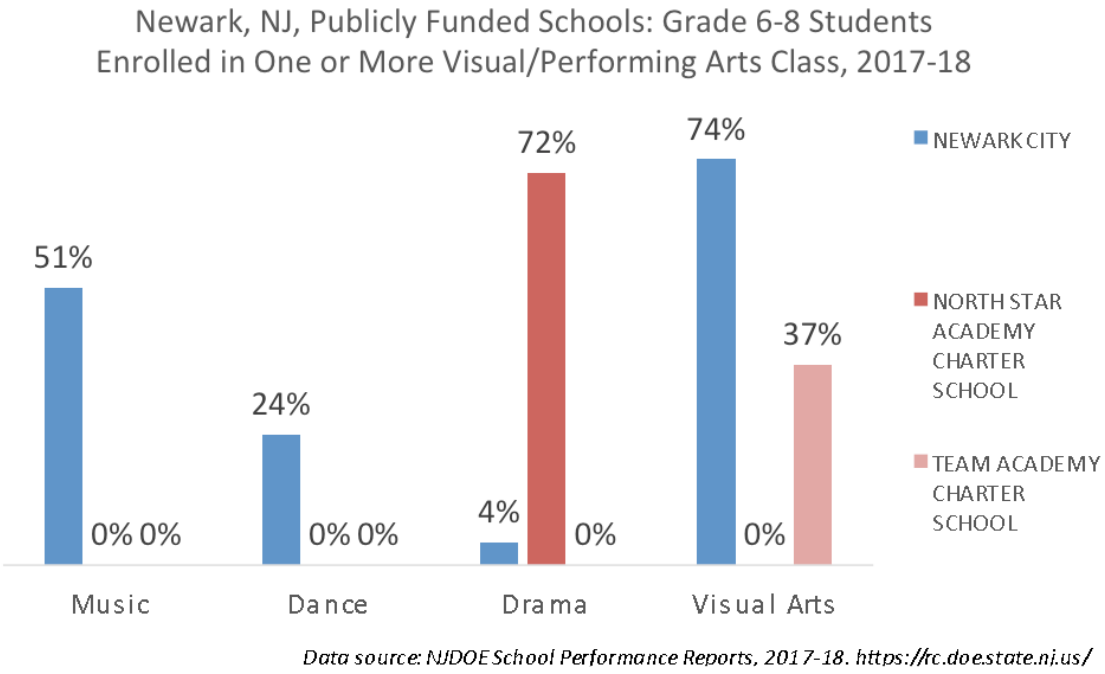

Figure 1 shows the opportunities for middle-school students in the arts across different Newark schools. Half of NPS students have classes in music; three-quarters have classes in the visual arts, and one-quarter study dance. In contrast, *none* of North Star/Uncommon's students receive any instruction in music, the visual arts, or dance. A little more than one-

third of TEAM/KIPP-NJ students take visual arts classes; they have no instruction in music, drama, or dance. This lack of arts instruction is likely to help keep the charters' instructional costs lower; it also allows for more curricular focus on tested subjects. The price, however, is an education that lacks the breadth that would come from including arts instruction.

VI. Review of the Validity of the Findings and Conclusions

The report makes a credible claim that it exploits the random component of Newark's school assignment system to find a substantial charter school effect for a subgroup of the city's students who expressed interest in enrolling in a charter. Less clear in the report, however, are the limitations of its findings; specifically:

- The effects the report finds can only be generalized to the population of students who apply to charter schools, which is likely quite different from the greater Newark population.
- The unique fiscal and staffing model found at the largest charters – which were found to have much larger effects on student achievement than other charters – relies on a relatively inexperienced and constantly churning teaching staff. There is no evidence that this model can be brought to scale if more students enroll in Newark's charters.
- The "no excuses" instructional model of the largest charters is based on a test-prep curriculum that excludes broad exposure to the arts, and likely other domains of learning.

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

Like similar studies, *Charter Schools in Newark: The Effect on Student Test Scores* credibly estimates the magnitude of a charter school effect; however, it sheds little light on *how* that effect is realized. The vague term "no excuses" does little to illuminate the unique staffing and fiscal structure, or the instructional and curricular narrowing, that likely contribute to the effect.

In addition, the limited generalizability of the findings, the questionable magnitude of the effect due to the nature of its two-stage model, and the ambiguous description of that effect as "large" are all reasons policymakers should exercise caution before using the findings of this report to justify the expansion of charter schools. Yes, Newark's charter schools appear to have a positive impact on test scores for a subgroup of students. But stakeholders should consider how those effects are realized, and the costs of those effects on the school system as a whole, before pursuing a program of charter school expansion.

Notes and References

1 See: Russakoff, D. (2016). *The prize: Who's in charge of America's schools?*. First Mariner Books edition. Boston, MA: Mariner Books/Houghton Mifflin Harcourt.

2 Winters, M.A. (2020). *Charter schools in Newark: The effect on student test scores*. New York, NY: Manhattan Institute. Retrieved January 15, 2020, from https://www.manhattan-institute.org/charter-school-effectiveness-newark-new-jersey

3 Throughout this review, I refer to all schools operated in Newark by KIPP-NJ as "TEAM/KIPP-NJ," and all schools operated by Uncommon as "North Star/Uncommon."

4 See the report, endnote #8.

5 Abdulkadiro□lu, A., Angrist, J.D., Narita, Y., Pathak, P.A (2017, September). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica 85*(5), 1373-1432.

6 Several studies using random school assignment to create a quasi-experiment are cited in the report's endnotes, including:

Abdulkadiro□lu, A., Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A. (2011, May.) Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics 126*(2), 699-74.

Cohodes, S.R.,Setren, E.M., Walters, C.R., Angrist, J.D., Pathak, P.A. (2013, October). *Charter school demand and effectiveness: A Boston update*. Boston Foundation and NewSchools Venture Fund.

Angrist, J.D., Pathak, P.A., & Walters,C.R. (2013, October). Explaining charter school effectiveness. *American Economic Journal: Applied Economics 5*(4), 1-27.

Angrist, J.D. et al. (2012). Who benefits from KIPP*? Journal of Policy Analysis and Management 31*(4), 837-60.

7 The use of 2SLS models, especially in research using charter school lotteries, is discussed extensively in:

Angrist, J.D., & Pischke, J.S. (2014). *Mastering metrics: The path from cause to effect*. Princeton, NJ: Princeton University Press. Angrist is a coauthor of several of the studies of charter schools cited in this report.

8 This 16% rate is for all Newark students, *including those enrolled in charters*. So the rate for students who did not apply to charters must be even higher. The same applies to the Limited English Proficient rates.

9 See footnote #5.

10 Kraft, M.A. (2018). *Interpreting effect sizes of education interventions*. Brown University Working Paper. Retrieved February 20, 2020, from https://scholar.harvard.edu/files/mkraft/files/kraft_2018_interpreting_effect_sizes.pdf

11 Examples include:
From the KIPP-NJ 2018 Annual Report: "In 2018, this vision is becoming a reality as Newark charter students have closed the proficiency gap with the state average." The report includes a comparison of proficiency rates on state tests to NJ state averages. https://kippnj.org/wp-content/themes/team/2018-annual-report-assets/pdf/2018-Annual-Report.pdf

An opinion piece in Newark's local newspaper included the following, credited to KIPP-NJ: "More about KIPP kids: Their elementary and high schools equal or outperform the average for the state of New Jersey, even though the students are much poorer; They close the achievement gap and surpass national averages

in reading and math by 8th grade." O'Connor, J. (2015, May 9). Beating Newark's odds, KIPP charter network is poised to expand. *The Star-Ledger*. Retrieved February 20, 2020, from https://www.nj.com/opinion/2015/05/beating_newarks_odds_kipp_charter_network_is_poise.html

The North Star/Uncommon Schools website compares the school's proficiency rates to those of South Orange and Verona public schools, two communities more affluent than Newark in Essex County, NJ. https://northstar.uncommonschools.org/results/

In an opinion piece from 2016, an administrator for Uncommon Schools compared North Star's proficiency rates to those of Millburn and Livingston High School, two communities more affluent than Newark in Essex County, NJ. Chigner, S. (2016, March 3). Myth of the hero teacher? What the New York Times gets wrong about schools overcoming poverty. *The 74*. Retrieved February 20, 2020, from https://www.the74million.org/article/myth-of-the-hero-teacher-what-the-new-york-times-gets-wrong-about-schools-overcoming-poverty/

12 Lemov, D. (2010). *Teach like a champion: 49 techniques that put students on the path to college*. San Francisco, CA: John Wiley & Sons.

13 Weber, M.A., & 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 February 20, 2020, from http://nepc.colorado.edu/thinktank/review-newark-reform

14 All data in this table is from the NJ Department of Education's data collection: https://www.nj.gov/education/data/. Staffing data is from NJDOE files made available through an Open Public Records Act (OPRA) request.

15 In addition, previous research has shown that the students with special needs enrolled in the Newark charters tend to have less-costly classifications than those in NPS schools. See: Weber, M. (2019).

16 The consequences of test-prep instruction is discussed here: Koretz, D. (2017). *The testing charade: Pretending to make schools bette*r. Chicago, IL: University of Chicago. The book specifically addresses claims made by Doug Lemov and Paul Bambrick-Santoyo, who both hold leadership positions at Uncommon Schools.